Mass, momentum and energy conserving (MaMEC) discretizations on general grids for the compressible Euler and shallow water equations

Hof, Bas van 't; Veldman, Arthur E.P.

Published in:
Journal of computational physics

DOI:
10.1016/j.jcp.2012.03.005

Publication date:
2012

Journal of Computational Physics 231 (2012) 4723–4744

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Mass, momentum and energy conserving (MaMEC) discretizations on general grids for the compressible Euler and shallow water equations

Bas van't Hof [a], Arthur E.P. Veldman [b,*]

[a] VORtech bv, P.O. Box 260, 2600 AG Delft, The Netherlands
[b] Institute for Mathematics and Computer Science, University of Groningen, P.O. Box 407, 9700 AK Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 3 October 2010
Received in revised form 6 February 2012
Accepted 12 March 2012
Available online 1 April 2012

Keywords:
CFD
Energy conserving discretization
Staggered grid
Shallow water equations
Compressible Euler equations

ABSTRACT

The paper explains a method by which discretizations of the continuity and momentum equations can be designed, such that they can be combined with an equation of state into a discrete energy equation. The resulting 'MaMEC' discretizations conserve mass, momentum as well as energy, although no explicit conservation law for the total energy is present. Essential ingredients are (i) discrete convection that leaves the discrete energy invariant, and (ii) discrete consistency between the thermodynamic terms. Of particular relevance is the way in which finite volume fluxes are related to nodal values. The method is an extension of existing methods based on skew-symmetry of discrete operators, because it allows arbitrary equations of state and a larger class of grids than earlier methods.

The method is first illustrated with a one-dimensional example on a highly stretched staggered grid, in which the MaMEC method calculates qualitatively correct results and a non-skew-symmetric finite volume method becomes unstable. A further example is a two-dimensional shallow water calculation on a rectilinear grid as well as on an unstructured grid. The conservation of mass, momentum and energy is checked, and losses are found negligible up to machine accuracy.

1. Introduction

Flow equations are considered consisting of a continuity and a momentum equation – expressing conservation of mass and momentum – plus a thermodynamic equation of state. No explicit equation for the evolution of total energy is present, but the latter can be derived analytically by combining the given equations. In absence of dissipative effects, total energy is then conserved, with an exchange between kinetic energy and thermodynamic potential energy. In the paper, discretizations of the flow equations will be discussed that similarly can be combined into a discrete conservation equation for total energy. The discussion will be held mainly for staggered computational grids, but it can also be applied to collocated grid arrangements. It is also extendable to formulations that include an internal energy equation, and will then likewise result in discrete conservation of total energy.

In this context, 'conservation' of a total quantity (like total mass, total momentum or total energy) does not mean that the total quantity does not change at all. The term 'conservation' is used to indicate that the total quantity changes *only* as a result of terms that change the quantity in the continuous model, of which the discretization is an approximation. Moreover, terms which change the total quantity only in a certain way in the continuous model (for instance, only create an increase or a decrease), should have the same property in the discretization. Hence, the discretization does not introduce any changes in the total amounts; only the appropriate physical terms do.

* Corresponding author.
E-mail addresses: bas@vortech.nl (B. van't Hof), a.e.p.veldman@rug.nl (A.E.P. Veldman).

0021-9991/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jcp.2012.03.005

In certain simulations, the conservation properties of the flow are very important for the correctness of the results. For instance, it is considered crucial in direct numerical simulation of turbulence (DNS), because the balance between energy dissipation and energy input determines the characteristics of the turbulent flow [5,18,31,36–38]. In shallow water simulations, both the momentum and energy balance are important in parts of the simulation domain where rapid fluctuations occur. In [26] a method is proposed that switches from a momentum conserving discretization to an energy conserving discretization, depending on the local flow characteristics. The method presented in the current paper is practical because the choice between momentum conservation and energy conservation is avoided: both are satisfied simultaneously.

In the papers mentioned thus far, the flow equations were mainly solved in a velocity-pressure formulation, either in conservative (divergence) form or in skew-symmetric form. An alternative is to formulate them in a rotational (vorticity) form. Especially in the literature regarding weather forecast, starting with Arakawa's seminal paper [1] (see also [16]), or shallow water equations, e.g. [2,24], this approach is followed. In this way it is possible to conserve energy in combination with enstrophy or helicity, e.g. [15], but the conservation of momentum is lost. This often is a serious drawback of the approach; see also the trade-off discussions on this issue in [18,19]. In a similar vain, in Hamiltonian particle-mesh methods the conservation of vorticity is advocated [6,9]. A comparison with such an approach would be interesting, as it is not yet known which quantity is the 'most relevant' to be preserved.

In many cases, finite volume methods are an effective way to construct conservative discretizations. However, there are two reasons why it is difficult to construct a fully conservative discretization method on a general (non-Cartesian) staggered grid. The main problem is that finite volume methods may be used to discretize the continuity and momentum equations, but in general it is not possible to combine the discrete continuity and momentum equations into a discrete energy equation. Already several decades ago, for incompressible flow [22] as well as compressible flow [7], it has been noted that the key ingredient for the construction of an energy-conserving discretization of the continuity and momentum equations is the preservation of the (skew)-symmetry properties of the differential operators in their discrete approximations. If these properties can be preserved, good results are obtained [8,10,14,31–33,36–38]. Also the related summation-by-parts schemes as introduced in [27] and the cosymmetry-preserving approach in [12] are designed to achieve this.

The other problem is that in a non-Cartesian staggered grid it is hard to conserve momentum with a finite volume method, because the discrete flow components each have a different direction. Yet, several attempts have been made in the literature to tackle the problem. A general framework is formed by the mimetic (or discrete calculus) methods developed at Los Alamos National Laboratory, e.g. [3,25]. Applications to unstructured staggered mesh schemes have been presented in [19–21,29].

The paper provides general guidelines for the construction of mass, momentum and energy conserving ('MaMEC') discretization schemes, also applicable to unstructured staggered grids. Of particular relevance is the way in which finite volume fluxes are related to nodal values. The method is first presented for a simple one-dimensional example of Euler flow in Section 3. Thereafter, more complex cases are discussed in Section 4. The order of accuracy and the size and shape of the discretization can still be chosen freely. An example of a shallow water simulation on a triangulated grid is presented in Section 5. While the current paper was under review, in [23] a similar method (also inspired by [38]) for the shallow water equations was presented.

A large variety of terms may be included in flow equations. For the sake of presentability, this paper discusses only the compressible Euler equations with a conservative force field and a general state equation. This choice of the terms included is motivated by the intended application to the shallow water equations. The shallow water equations are mathematically equivalent to the compressible Euler equations with a specific choice of the state equation. The conservative force field is necessary if water flow over a non-flat bottom profile is to be described.

The paper describes the construction of discretizations inside the computational domain; the discretization of boundaries is outside its scope. The examples presented avoid the problem of boundary condition discretization by applying periodic boundary conditions. Careful analysis of the desired properties of boundary conditions is necessary to obtain discretizations for calculations on domains with actual boundaries; see e.g. [4] for an immersed boundary (cut cell) treatment with application in [13,34].

Complementary to the above structure-preserving attention paid to the space discretization, also the time integration can be reconsidered. In the recent literature several time integration ideas have been presented to preserve discrete energy and/or other invariants, e.g. [28,11,17]. This aspect has been left outside the scope of the present paper. In order for the time integration errors not to 'pollute' the discretization study, time steps have been taken exaggeratingly small.

Remark. The above approach of preserving certain selected quantities during discretization, can also be applied during the process of modeling. E.g. in [35,30] turbulence models are presented that preserve energy (in the absence of diffusion), enstrophy and helicity.

2. Flow equations

2.1. Euler equations

In this section, we present the Euler equations for compressible or incompressible flow, with a conservative force field whose potential is ϕ, given by

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \boldsymbol{v} = 0 \text{ continuity} \tag{1}$$

$$\frac{\partial}{\partial t} \rho \boldsymbol{v} + \nabla \cdot (\rho \boldsymbol{v} \otimes \boldsymbol{v}) + \nabla p = -\rho \nabla \phi \text{ momentum} \tag{2}$$

$$\rho = R(p) \text{ equation of state,} \tag{3}$$

where the following notations are used:

ρ	Density [kg/m^d]
d	Dimension 1, 2 or 3
$\boldsymbol{v}$	Flow velocity [m/s]
p	Pressure [N/m^{d-1}]
ϕ	Potential function [m^2/s^2]

The equations are written in divergence form, so the conservation laws for mass and momentum can be recognized. The term on the right hand side in (2) shows that the conservative force field is the only term which may change the total momentum.

The continuity and momentum equations can be combined into the energy equation. For this purpose, we introduce the *energy density* $e := e_{kin} + e_{pot}$, which is the sum of the kinetic and potential energy densities, given by

$$e_{kin} = \frac{1}{2}|\boldsymbol{v}|^2, \quad e_{pot} = \phi + \Phi - \frac{p}{\rho},$$

where the *energy transport potential* Φ is given by

$$\Phi := \int^p \frac{1}{R(s)}\, \mathrm{d}s.$$

It is easy to see that in the incompressible case, where the density function R is constant, the potential energy is equal to the potential ϕ. At several places below, the following derivatives of Φ will be needed:

$$\frac{\partial \Phi}{\partial t} = \frac{1}{R(p)} \frac{\partial p}{\partial t} = \frac{1}{\rho} \frac{\partial p}{\partial t} \quad \text{and} \quad \nabla \Phi = \frac{1}{\rho} \nabla p. \tag{4}$$

An energy equation can be derived by combining the continuity and momentum equations in the following way

$$\boldsymbol{v} \cdot \text{ Eq. (2)} + \left(\Phi + \phi - \frac{1}{2}|\boldsymbol{v}|^2\right) \text{ Eq. (1)}, \tag{5}$$

which results in (see Appendix A.2)

$$\frac{\partial \rho e}{\partial t} + \nabla \cdot \left(\rho \boldsymbol{v} \left(e + \frac{p}{\rho}\right)\right) = 0. \tag{6}$$

A divergence form exists, showing that there are no terms in the equations that change the total amount of energy.

This paper discusses the way in which discretizations of the continuity and momentum equations can be designed such that they can be combined into a meaningful discrete energy equation with a discrete divergence form similar to the continuous divergence form in (6). Essential ingredients are the numerical discretization of the convective terms in the momentum equation (2) and the numerical treatment of the thermodynamic relation (4). Of particular relevance is the way in which finite volume fluxes are related to nodal values.

The method can be employed regardless of the specific equation of state which is used, including the incompressible equation with a constant density and ideal gasses with a linear equation of state. The examples in this paper are based on the shallow water equations, which we show below to be a special case of the compressible Euler equations, and in which certain terms can be simplified.

2.2. Shallow water equations

The shallow water equations are a two-dimensional model of water motion. The vertical dimension is removed from the system by the hydrostatic pressure approximation, and the flow velocity is approximated by the depth-averaged flow velocity. The following notations are used:

ρ_w	Density of water 1000 kg/m^3
z_b	z-Coordinate of the bottom [m]
ζ	z-Coordinate of the water surface [m]

(*continued on next page*)

$H = \zeta - z_b$	Water height [m]
g	Gravitational acceleration 9.8 m/s^2
$H_e = \zeta + \lvert\boldsymbol{v}\rvert^2/2g$	Energy head [m]

These shallow water variables are related to the variables introduced in Section 2.1 by

$\rho = \rho_w H$	Density [kg/m^2]
$p = g\rho_w H^2/2$	Depth-integrated pressure [kg/s^2] = [N/m]
$\Phi = gH$	Energy transport function [m^2/s^2] = [J/kg]
$\phi = g\, z_b$	Potential field [m^2/s^2] = [J/kg]
$e_{pot} = g(\zeta + z_b)/2$	Potential energy density [m^2/s^2] = [J/kg]

Using these notations, the flow equations become

$$\begin{aligned}
&\frac{\partial H}{\partial t} + \nabla \cdot H\boldsymbol{v} = 0 \text{ continuity}\\
&\frac{\partial}{\partial t} H\boldsymbol{v} + \nabla \cdot (H\boldsymbol{v} \otimes \boldsymbol{v}) + g\nabla H^2/2 = -gH\nabla z_b \text{ momentum}\\
&\frac{\partial}{\partial t}\left(\frac{H|\boldsymbol{v}|^2 + g\zeta^2}{2}\right) + \nabla \cdot (gH\boldsymbol{v}H_e) = 0 \text{ energy}
\end{aligned} \tag{7}$$

Again, the energy equation (7) follows from a combination of the equations for mass and momentum similar to (5).

3. MaMEC discretization on a non-uniform 1D grid

In this section, a discretization is presented for the one-dimensional compressible Euler equations that, next to mass and momentum, also conserves energy. In this, relatively simple, example all details of the approach can be demonstrated. In particular, all required (and sometimes subtle) relations between the various ingredients of the discretization become visible. At the end of the section, this discretization is applied to the shallow water equations (7).

The computational grid, illustrated in Fig. 1, consists of pressure and velocity points. The pressure and velocity points are organized as in a standard Arakawa-C grid [16]. The positions of grid points (pressure points) and cell faces (velocity points) are allowed to be arbitrary. In order to achieve the envisaged MaMEC properties, it is not necessary that one set of points should lie halfway the other set of points. This leaves room to let accuracy decide which choice to make.

3.1. Conservative discretization formulas

As a first example, the MaMEC discretization is described for the one-dimensional Euler equations. The conservation properties are demonstrated by writing the mass, momentum and energy equations in their finite volume form. The essential aspects of the general method for the construction of MaMEC discretizations are pointed out in this section; the general derivation follows in the following sections of this chapter.

3.1.1. Discrete continuity equation

We use the following finite volume type discrete continuity equation applied to the cell around grid point x_i:

$$\frac{\partial \rho_i}{\partial t} + \frac{q_{i+1/2} - q_{i-1/2}}{\Delta x_i} = 0, \tag{8}$$

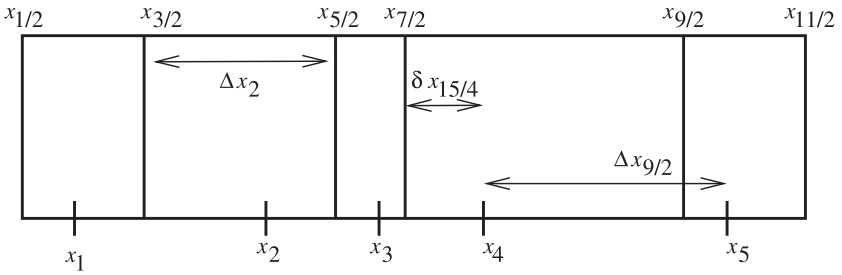

Fig. 1. 1D-staggered grid: the pressure point i lies between the velocity points $i - 1/2$ and $i + 1/2$. The grid distances are such that $\Delta x_i = \delta x_{i-1/4} + \delta x_{i+1/4}$.

where the discharges $\mathtt{q}$ are defined by

$$\mathtt{q}_{i+1/2} \;:=\; \mathtt{v}_{i+1/2}\;\hat{\rho}_{i+1/2}, \tag{9}$$

with the discharge densities $\hat{\rho}$ given by

$$\hat{\rho}_{i+1/2} \;:=\; \frac{\mathtt{p}_{i+1}-\mathtt{p}_i}{\Phi_{i+1}-\Phi_i}. \tag{10}$$

The reason for choosing the discharge density $\hat{\rho}$ according to (10) is the product rule

$$\mathtt{v}_{i+1/2}\frac{\mathtt{p}_{i+1}-\mathtt{p}_i}{\Delta x_{i+1/2}} = \mathtt{v}_{i+1/2}\;\hat{\rho}_{i+1/2}\frac{\Phi_{i+1}-\Phi_i}{\Delta x_{i+1/2}} = \mathtt{q}_{i+1/2}\frac{\Phi_{i+1}-\Phi_i}{\Delta x_{i+1/2}}, \tag{11}$$

similar to (4)

$$v\frac{\partial p}{\partial x} = v\rho\;\frac{\partial \Phi}{\partial x} = q\frac{\partial \Phi}{\partial x}.$$

Note that the calculation of $\hat{\rho}$ according to (10) need not fail where the denominator is (near) zero, since the product rule (11) is easily satisfied in that case. In the case of the shallow water equations, the discharge water height $H_{i+1/2}$ (which plays the role of the discharge density $\hat{\rho}_{i+1/2}$ in the Euler equations) is simply the average of the water heights H_i and H_{i+1}. This simple 'half-half' (equal-weight) average also has to be used on non-uniform grids, where one is tempted to apply linear interpolation; see the example in Sections 3.2 and 3.3. The discharge density $\hat{\rho}$ is used only in the calculation of the discharge. Because of the staggering of the grid it is defined at cell faces, whereas the 'normal' density is defined in cell centers.

3.1.2. Discrete momentum equation

The momentum equation, governing the evolution of $\rho\mathtt{v}$, is discretized at cell faces. Hence it also requires a cell-faced density $\rho_{i+1/2}$, which is not necessarily equal to the above-defined discharge density $\hat{\rho}_{i+1/2}$. The need for two different cell-faced densities can be explained by noting that the local discharges and the local momentum, though both equal to $\rho\boldsymbol{v}$, are essentially different quantities: discharges are cell-face averages, and the local amounts of momentum are cell-volume averages. The cell-faced density $\rho_{i+1/2}$ to be used in the calculation of the local momentum is defined as

$$\rho_{i+1/2} \;:=\; \frac{\delta x_{i+1/4}\;\rho_i + \delta x_{i+3/4}\;\rho_{i+1}}{\Delta x_{i+1/2}}, \tag{12}$$

where the cell-volume averaging is expressed in the choice of the interpolation weights (a motivation of this choice follows below). Whether we integrate the cell-centered densities ρ_i or the cell-faced densities $\rho_{i+1/2}$, the same total amount of mass is found with this choice for the cell-faced density (ignoring boundary effects):

$$\Sigma_i \Delta x_{i+1/2}\;\rho_{i+1/2} = \Sigma_i(\delta x_{1+1/4}\;\rho_i + \delta x_{i+3/4}\;\rho_{i+1}) = \Sigma_j(\delta x_{j-1/4} + \delta x_{j+1/4})\rho_j = \Sigma_j \Delta x_j\;\rho_j. \tag{13}$$

Also, as we will see below, the weights chosen in (12) make it possible to combine the continuity equation (8) with the definition of the cell-faced density $\rho_{i+1/2}$ into a cell-faced continuity equation of the following form:

$$\frac{\partial \rho_{i+1/2}}{\partial t} + \frac{\mathtt{q}_{i+1}-\mathtt{q}_i}{\Delta x_{i+1/2}} = 0, \tag{14}$$

in which the discharges $\mathtt{q}_i$ at full indices, i.e. in cell centers, are defined as

$$\mathtt{q}_i \;:=\; \frac{\delta x_{i-1/4}\;\mathtt{q}_{i+1/2} + \delta x_{i+1/4}\;\mathtt{q}_{i-1/2}}{\Delta x_i}. \tag{15}$$

Note that (15) corresponds to a linear interpolation of the discharges $\mathtt{q}_{i+1/2}$ in (9) from cell-faces to cell-centers. It is stressed that (14) holds irrespective of the precise choice of $\mathtt{q}_{i+1/2}$; only (15) has to hold. The form of (15) is fully fixed once the choice (12) has been made. The proof of (14), given in Appendix A.1, reveals how the weights in (12) and (15) have a subtle relation.

Next, we are ready to introduce the following discrete momentum equation in finite volume form:

$$\frac{\partial}{\partial t}(\rho_{i+1/2}\mathtt{v}_{i+1/2}) + \frac{\mathtt{FV}_{\,i+1}-\mathtt{FV}_{\,i}}{\Delta x_{i+1/2}} = -\hat{\rho}_{i+1/2}\frac{\phi_{i+1}-\phi_i}{\Delta x_{i+1/2}}, \tag{16}$$

where the discrete momentum flux $\mathtt{FV} \doteq qv + p$ has been chosen as

$$\mathtt{FV}_{\,i} \;:=\; \mathtt{q}_i\frac{\mathtt{v}_{i+1/2}+\mathtt{v}_{i-1/2}}{2} + \mathtt{p}_i, \tag{17}$$

with $\mathtt{q}$ from (15). Here the cell-faced velocities are combined into a cell-centered value by simple equal-weight averaging, i.e. without using geometry information. An important property of this convective discretization is that it generates a skew-symmetric convective contribution to the coefficient matrix. This will turn essential later on (see Section 4.1.5), The symmetry of (17) with respect to $\mathtt{v}_{i+1/2}$ and $\mathtt{v}_{i-1/2}$ is a necessary condition hereto. The relevance of this property will become clear when the discrete energy equation is formed.

Note that only the right-hand side of the momentum equation (16), i.e. the conservative force field, causes a change in the total discrete momentum. It has to be consistent with the thermodynamic terms, hence the appearance of the discharge density $\hat{\rho}$.

3.1.3. Discrete energy equation

The discrete continuity and momentum equations have been carefully constructed so that they can be combined into a discrete energy equation in a finite volume form, which expresses discrete energy conservation. In particular, the choices (10) and (12) of the two cell-faced density values and the choice for the advection term in the momentum flux definition (17) are essential for discrete energy conservation. A general explanation of the method for the construction of such a momentum equation is given later in Section 4.1. Here, first a one-dimensional version is explained in detail.

Starting point is the discrete choice for the energy $\rho e = \rho(e_{kin} + e_{pot})$, in the grid point $x_{i+1/2}$. It is defined, similar to (12), as

$$(\rho e)_{i+1/2} \;:=\; \frac{1}{2}\rho_{i+1/2}\mathrm{v}_{i+1/2}^2 + \frac{\delta x_{i+1/4}}{\Delta x_{i+1/2}}\rho_i e_{pot,i} + \frac{\delta x_{i+3/4}}{\Delta x_{i+1/2}}\rho_{i+1}e_{pot,i+1} \doteq (\rho e_{kin})_{i+1/2} + (\rho e_{pot})_{i+1/2}. \tag{18}$$

After some arithmetic, presented in Appendix A.3, the time differentiation of (18) results in a discrete energy equation which has the finite volume form

$$\frac{\partial(\rho e)_{i+1/2}}{\partial t} + \frac{\mathtt{FE}_{\,i+1} - \mathtt{FE}_{\,i}}{\Delta x_{i+1/2}} = 0, \tag{19}$$

where the discrete version of the energy flux $\mathtt{FE} \doteq q\,(e + p/\rho) = q\,\left(\frac{1}{2}|v|^2 + \Phi + \phi\right)$ in the grid point x_i is given by

$$\mathtt{FE}_{\,i} \;:=\; \mathrm{q}_i\left(\frac{\mathrm{v}_{i+1/2}\mathrm{v}_{i-1/2}}{2} + \Phi_i + \phi_i\right). \tag{20}$$

Observe that during the derivation of this expression use has been made of the properties (9) and (11) related to the discharges, and of the carefully chosen interpolations (15), (17) and (18). We do advise the interested reader to carefully walk through the steps of this derivation, as worked out in Appendix A.3, since in the sequel we will present the same reasoning in a more abstract setting.

The finite volume form of the discrete energy equation ensures energy conservation, both globally and locally. As explained in Appendix B (Lemma 2), skew symmetry of the discrete convection operator is a necessary condition to obtain such a divergence form. Finite volume (i.e. local conservation) forms of the continuity, momentum and energy equations similar to the ones presented in this section exist in all MaMEC discretizations discussed in this paper. Only in this section are they given explicitly.

As discussed already in the introduction, many discretizations of the continuity and momentum equations cannot be combined into a finite volume energy equation, and therefore introduce numerical energy losses or gains. The latter may cause instabilities, as we will show in Section 2.2. First, we shall introduce such a 'lossy' discretization in Section 3.2, so we can compare the 'lossy' finite volume results to the MaMEC results in Section 3.3.

3.2. Finite volume discretization with linear interpolation

The conservative MaMEC discretization formulas from the previous section are now compared to the usual second-order central finite volume discretization. Instead of having two interpolations for (9) and (10), this discretization only uses the linear interpolation $\bar{\rho}$:

$$\bar{\rho}_{i+1/2} \;:=\; \rho_i + \frac{\rho_{i+1} - \rho_i}{x_{i+1} - x_i}\,(x_{i+1/2} - x_i). \tag{21}$$

The same linear interpolation is used to calculate the interpolated velocity $\bar{\mathrm{v}}$:

$$\bar{\mathrm{v}}_i \;:=\; \mathrm{v}_{i-1/2} + \frac{\mathrm{v}_{i+1/2} - \mathrm{v}_{i-1/2}}{x_{i+1/2} - x_{i-1/2}}\,(x_i - x_{i-1/2}). \tag{22}$$

It is stressed that on a non-uniform grid, this interpolation differs from the 'half-half' (equal-weight) average as in (17) which is required to obtain a MaMEC discretization; see Section 3.1.1. Thus, it results in a different value for the discharge $\mathrm{q}_{i+1/2}$ and hence in a different continuity equation.

The discrete momentum equation used is

$$\frac{\partial}{\partial t}(\bar{\rho}_{i+1/2}\mathrm{v}_{i+1/2}) + \frac{\mathtt{FV}_{\,i+1} - \mathtt{FV}_{\,i}}{x_{i+1} - x_i} + \frac{\Phi_{i+1} + \phi_{i+1} - \Phi_i - \phi_i}{x_{i+1} - x_i} = 0, \tag{23}$$

where the discrete momentum flux $\mathtt{FV} \doteq qv + p$ is chosen as $\mathtt{FV}_{\,i} = \rho_i\bar{\mathrm{v}}_i^2 + \mathrm{p}_i$. In irregular grids, the coefficient of v_i is not always zero, hence the convective operator is not always skew symmetric. Thus, the given momentum (23) and continuity (8) equations cannot be combined into a finite volume form of the energy equation (cf. Lemma 2 in Appendix B which states that skew symmetry is a necessary condition). As a consequence, they may introduce numerical energy losses (or gains), as will be seen in the numerical example below.

3.3. Results for the shallow water equations

The conservative discretizations from the previous two sections (MaMEC from Section 3 and finite volume with linear interpolation from Section 3.2) are now compared for the shallow water equations in a one-dimensional simulation. Though it is possible to construct special MaMEC time integrators [11,17,28], we opted to perform the time integration with a 'standard' Runge-Kutta method, however with very small time steps. This reduces the effects of the time integration method.

The domain is a canal, 100 meters in length, with a bottom profile given by the sum of two sine functions and periodic boundary conditions. A non-uniform grid is used, as can be seen in Fig. 2. For the initial solution, we take the exact solution of the stationary shallow water equations, as is shown in Fig. 2. Due to discretization errors, the discrete solutions are not stationary. The small waves which appear in the canal never diminish, because there are no dissipative terms in the model at all, neither continuous (analytical) nor discrete (numerical).

The changes in total momentum and energy in the two calculations are compared in Fig. 3 and the final solutions in Fig. 4. Both methods predict that the total momentum changes in time. This is due to the non-flat bottom, which introduces a source term in the momentum equation. The MaMEC method predicts that the total momentum increases and decreases in a non-periodic way. The total momentum never changes more than 0.1%. The finite volume method with linear interpolation (21), (22) predicts a steadily decreasing total momentum, which physically is incorrect. This is due to the incompatible treatment of the bottom-related source terms in (23).

The second graph in Fig. 3 shows the change in energy. The MaMEC method predicts an energy which remains constant within the accuracy of the time integration. The method with linear interpolation predicts an energy gain during the first 497 simulated seconds, after which the solution becomes unstable. The effect of the instability of the non-skew-symmetric linearly-interpolating finite volume method is shown in Fig. 4: wiggles have developed in the water level and the flow velocity field.

In this case, the non-skew-symmetric method breaks down under its energy losses, while the symmetry-preserving MaMEC method produces accurate results: the solution remains close to the initial conditions, which were after all a stationary solution of the shallow water equations. Note that the example has been especially designed to trigger failure for the non-skew-symmetric method, in order to demonstrate the qualitative properties of a MaMEC discretization. Even for this highly irregular mesh, the amount of energy is under control (constant within the accuracy of the time integration), which is the reason why the MaMEC calculation remains stable. It should be remarked, that more of the mimetic methods mentioned in Section 1 share this property.

4. Construction of conservative methods on more general classes of grids

In this section the one-dimensional discretization formulas from the previous section will be extended to two dimensions and to more general computational grids. To begin with, Section 4.1 presents the discretization on rectilinear (Cartesian) grids with a lower (second) order gradient discretization. Most of the concepts of MaMEC (mass, momentum and energy conserving) discretizations, as well as most of the necessary notations, are introduced in this section. In its last subsection, the abstract presentation is elucidated by relating it to the one-dimensional situation described in Section 3. Finally, Section 4.2

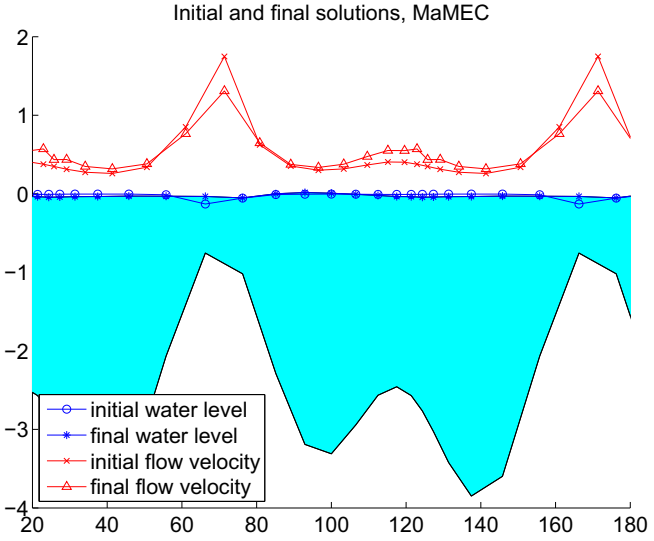

Fig. 2. The initial solution for the 1D test corresponds to the exact stationary solution of the shallow water equations. The markers show the location of the grid points in the non-uniform grid, which deliberately has been chosen highly distorted to create a numerical challenge.

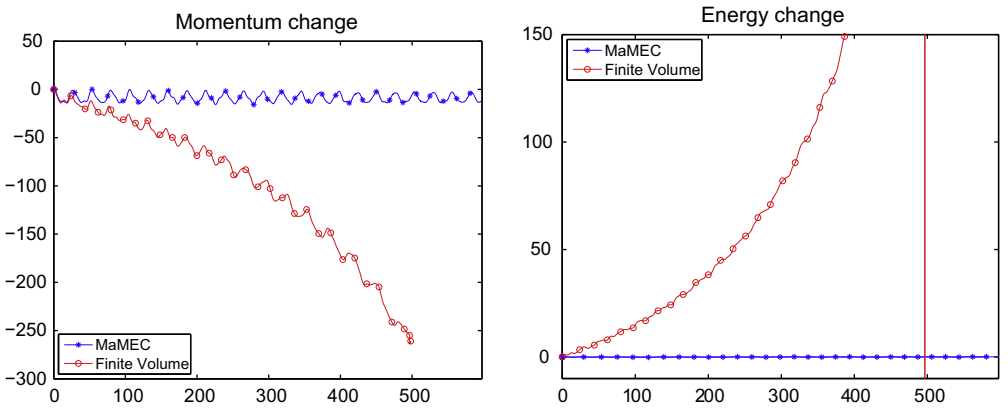

Fig. 3. The losses of momentum and energy over time. Due to the non-flat bottom, the total momentum is not constant in time. Unlike the MaMEC method, however, the finite volume method with linear interpolation predicts a steadily decreasing momentum and increasing energy, until the end of the simulation when the solution becomes unstable.

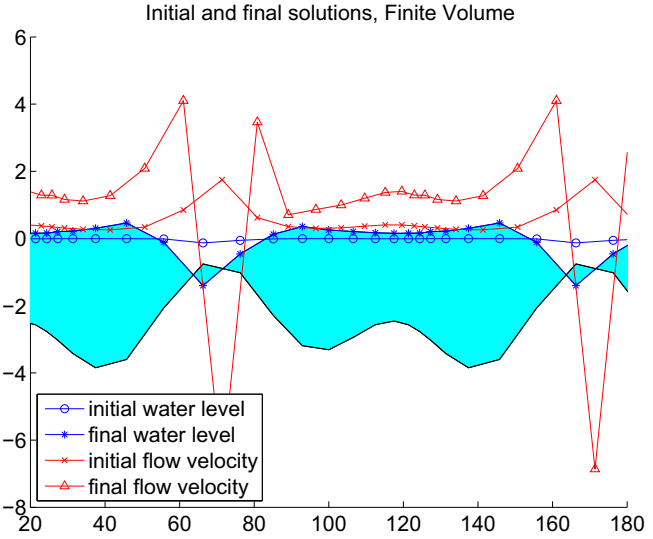

Fig. 4. Water level and velocities in the final solution (t = 497 s). The solution of the non-skew-symmetric finite volume method has developed wiggles as a result of energy losses: the water level has gone below the bottom depth.

generalizes the approach for discretizations with a (more complicated) larger gradient stencil which can be needed on unstructured grids; an example is included.

4.1. MaMEC discretizations on rectilinear grids with a two-point gradient stencil

In this section, the discretization presented in Section 3 is generalized for the two- or three-dimensional Euler equations. The discussion of the discretizations will be more abstract than the discussion in Section 3, because notations are introduced which are needed to construct more general classes of grids. The construction of the discretizations in this section involves all the essential elements of the MaMEC discretization technique: the construction of discrete chain rules and the preservation of the continuous operators' (skew-)symmetry in their discrete approximations.

Notation

To distinguish between discrete flow variables (grid functions) defined in separate positions of the grid, some notation is introduced. It is assumed that all functions (and their derivatives) vanish at the boundaries, in order not to be bothered with boundary conditions.

- The grid functions defined in density points belong to the space $\mathcal{V}_\rho$. The constant grid function with unit entries everywhere is denoted $\mathbf{1}_\rho \in \mathcal{V}_\rho$. A discrete inner product is defined by (note that only real-valued functions are considered)

$$\langle\langle \phi, \psi \rangle\rangle_\rho \;:=\; \Sigma_i \mathtt{w}_{\rho i}\, \phi_i\, \psi_i \quad (\phi, \psi \in \mathcal{V}_\rho), \tag{24}$$

  where the summation runs over all participating grid points and $\mathtt{w}_\rho$ denotes the discrete integration weight: say, the size of the corresponding control volume.
- Grid functions defined in the locations where the x-velocity component v_x is defined make up the space $\mathcal{V}_x$. Its discrete inner product is denoted

$$\langle\langle u, v \rangle\rangle_x \;:=\; \Sigma_i \mathtt{w}_{xi}\, u_i\, v_i \quad (u, v \in \mathcal{V}_x), \tag{25}$$

  with $\mathtt{w}_x$ the corresponding integration weights. The unit operator in this space is denoted $\mathtt{I}_x : \mathcal{V}_x \to \mathcal{V}_x$, whereas the constant function with unit entries everywhere is denoted $\mathbf{1}_x \in \mathcal{V}_x$. Similar definitions hold for the other two velocity components. Finally, the discrete velocity vector $\mathbf{v} = (v_x, v_y, v_z)$ lives in the space $\mathcal{V}_3 \equiv \mathcal{V}_x \times \mathcal{V}_y \times \mathcal{V}_z$.

4.1.1. Discrete approximations for integrals and total amounts

This paper is concerned with the conservation of mass, momentum and energy. As was explained in Section 1, the term 'conservation' refers to the changes in the total amounts of mass, momentum and energy. Compact notations for the discrete approximations of these total amounts are introduced in this section. The following total amounts are essential in the MaMEC discretization for compressible fluid flow:

Total mass M_ρ The total mass is the integral of the density ρ:

$$M_\rho = \int_\Omega \rho(\mathbf{x})\, \mathrm{d}\Omega.$$

The discrete approximations of the densities at a chosen set of 'density points' are combined into the vector $\rho \in \mathcal{V}_\rho$. The total mass is approximated with help of the inner product (24) as defined in $\mathcal{V}_\rho$, in which the size of the control volumes has been incorporated already:

$$M_\rho \approx \langle\langle \mathbf{1}_\rho, \rho \rangle\rangle_\rho.$$

Choosing the cell-centers as 'density points' and the control volume areas as integration weights results in a second-order accuracy for the approximated total mass. A higher order of accuracy may be obtained by choosing different weights.

Total momentum $\mathbf{M}_v := (M_x, M_y, M_z)$. The total momentum is the integral of the local momentum vector $\rho\boldsymbol{v}$:

$$\mathtt{M}_v = \int_\Omega \rho \boldsymbol{v}(\mathbf{x})\, \mathrm{d}\Omega.$$

The calculation of local momentum requires the interpolation of the mass using the interpolation matrices $\mathtt{S}_x : \mathcal{V}_\rho \to \mathcal{V}_x$, $\mathtt{S}_y : \mathcal{V}_\rho \to \mathcal{V}_y$ and $\mathtt{S}_z : \mathcal{V}_\rho \to \mathcal{V}_z$, because each of the velocity components is approximated on its own grid, which is not the same as the grid used for the densities. The interpolations are used to obtain the following approximations for the total momentum components M_x, M_y and M_z:

$$M_x \approx \langle\langle \mathbf{1}_x, (S_x\rho)\; v_x \rangle\rangle_x, \quad M_y \approx \langle\langle \mathbf{1}_y, (S_y\rho)\; v_y \rangle\rangle_y, \quad M_z \approx \langle\langle \mathbf{1}_z, (S_z\rho)\; v_z \rangle\rangle_z, \tag{26}$$

where $\mathtt{v}_x \in \mathcal{V}_x$, etc., are vectors holding the approximated flow velocity components. Incorporated already in the inner products as defined in (25) are the integration weights $\mathtt{w}_x \in \mathcal{V}_x$, etc., used for the approximate integration of staggered fields (Fig. 5).

Analogous to (13), the interpolations are chosen to satisfy a compatibility relation such that the integral of all the (interpolated) density fields is the same:

$$\langle\langle \mathbf{1}_\rho, \rho \rangle\rangle_\rho = \langle\langle \mathbf{1}_x, \mathtt{S}_x\rho \rangle\rangle_x = \langle\langle \mathbf{1}_y, \mathtt{S}_y\rho \rangle\rangle_y = \langle\langle \mathbf{1}_z, \mathtt{S}_z\rho \rangle\rangle_z \quad \forall \rho \in \mathcal{V}_\rho. \tag{27}$$

Total kinetic energy E_{kin}. The kinetic energy E_{kin} can be split up into the contributions $E_{kin,x}$, $E_{kin,y}$ and $E_{kin,z}$ for each of the velocity components, with each component its own discrete integration volume; see Fig. 5. The contribution of the x-component of the velocity is given by

$$E_{kin,x} = \frac{1}{2} \int_\Omega \rho(\mathrm{x})\, v_x(\mathrm{x})^2\, \mathrm{d}\Omega \approx \frac{1}{2} \langle\langle \mathtt{v}_x,\; (\mathtt{S}_x\rho)\; \mathtt{v}_x \rangle\rangle_x. \tag{28}$$

Note that all factors in the right-hand side belong to $\mathcal{V}_x$. The other contributions can be constructed analogously.

Total potential energy E_{pot}.

$$E_{pot} = \int_\Omega \{\rho(\Phi + \phi) - p\}\, \mathrm{d}\Omega \approx \langle\langle \mathbf{1}_\rho, \rho(\Phi + \phi) - \mathtt{p} \rangle\rangle_\rho, \tag{29}$$

where $\mathtt{p}$ and Φ are the vectors in $\mathcal{V}_\rho$ containing the pressure and the energy transport potential, introduced in Section 2:

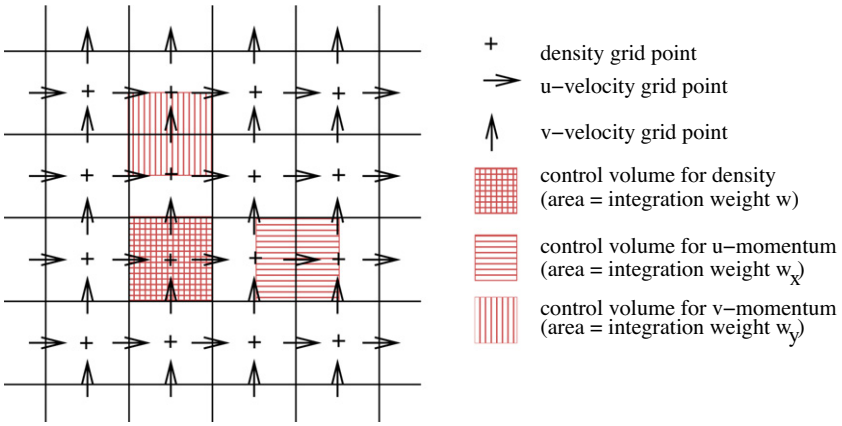

Fig. 5. Integration volumes for density and momentum.

$$\rho = R(\mathtt{p}), \quad \Phi = \Phi(\mathtt{p}).$$

Total energy.

$$E = E_{pot} + E_{kin}.$$

Note that the two contributions (28) and (29) are integrated on different grids.

4.1.2. Discrete continuity and momentum equations

Mass. The discrete continuity equation is given in finite volume form as

$$\frac{\partial \rho}{\partial t} + \mathtt{DIVX}\ \mathtt{q}_x + \mathtt{DIVY}\ \mathtt{q}_y + \mathtt{DIVZ}\ \mathtt{q}_z = 0, \tag{30}$$

where $\mathtt{DIVX} : \mathcal{V}_x \to \mathcal{V}_\rho$, etc., are the component-wise contributions to the divergence operator, and where the vectors $\mathtt{q}_x \in \mathcal{V}_x$, etc., represent the components of the discharge vector field, $\mathtt{q} = (\mathtt{q}_x, \mathtt{q}_y, \mathtt{q}_z)$, analytically given by $q = \rho \boldsymbol{v}$. Its discrete approximation will be discussed in this section. Sometimes, for shorthand, we will abbreviate

$$\mathtt{DIV}\ \mathtt{q} \ := \ \mathtt{DIVX}\ \mathtt{q}_x + \mathtt{DIVY}\ \mathtt{q}_y + \mathtt{DIVZ}\ \mathtt{q}_z. \tag{31}$$

The definitions of the discrete partial derivatives and the discrete discharge are presented in Section 4.1.4. In order to obtain discrete energy conservation, the partial derivatives have to satisfy a generalized version of the product rule (11) for the pressure and potential energy function. It will be seen in Section 4.1.4 that this leads to an approximation similar to (9).

Mass conservation (no change in time for M_ρ, i.e. $\partial M_\rho / \partial t = 0$) is guaranteed when integrals of partial derivatives are zero:

$$\langle\langle \mathbf{1}_\rho, \mathtt{DIVX}\ \mathtt{v}_x \rangle\rangle_\rho = 0 \quad \forall \mathtt{v}_x \in \mathcal{V}_x, \ \text{etc.} \tag{32}$$

In a finite volume context, starting from a divergence form of the equations and assuming homogeneous or periodic boundary conditions, this property follows immediately from telescoping the discrete divergence.

Momentum. The equation for momentum in x direction will have the following form

$$\frac{\partial}{\partial t}\{(\mathtt{S}_x \rho)\mathtt{v}_x\} + \mathtt{ADV}_x\ \mathtt{v}_x + \mathtt{GRADX}\ \mathtt{p} + \hat{\rho}_x\ \mathtt{GRADX}\ \phi = 0. \tag{33}$$

Here, $\hat{\rho}_x$ denotes a discharge density corresponding with the x-component of the gradient (i.e. $\mathtt{GRADX} : \mathcal{V}_\rho \to \mathcal{V}_x$). The equations for momentum in the other directions are equivalent. The following operators and vectors have been introduced in (33):

- $\mathtt{ADV}_x : \mathcal{V}_x \to \mathcal{V}_x$: the x-component of the discrete advection operator. It should approximate the following continuous advection operator:

$$\mathtt{ADV}\ v_x \doteq \nabla \cdot (\mathtt{q}\ v_x), \tag{34}$$

where v_x denotes the x-component of the continuous velocity field. The way in which the discrete advection operator approximates the continuous operator, and the order of accuracy, is not the subject of this paper. This paper describes certain properties that the discrete operators should have in order to be mass, momentum and energy conserving. The accuracy de-

pends on the specific choice for the operators and, in principle, any given order of accuracy may be achieved. An example of a fourth order MaMEC discretization on a staggered Cartesian grid is presented in [38].

- $\texttt{GRADX} : \mathcal{V}_\rho \to \mathcal{V}_x$: the x-component of the discrete gradient operator.

Momentum conservation Combining (26) and (33), the time derivative of the total (integrated) momentum is given by

$$\frac{\partial M_x}{\partial t} = -\langle\langle \mathbf{1}_x, \texttt{ADV}_x\ \texttt{v}_x + \texttt{GRADX p} + \hat{\rho}_x \texttt{GRADX}\ \phi\rangle\rangle_x \tag{35}$$

Momentum conservation is guaranteed when both the advection and the gradient operators have zero integral, i.e. are of divergence form (again homogeneous or periodic boundary conditions are assumed):

$$\langle\langle \mathbf{1}_x, \texttt{ADV}_x\ \texttt{v}_x\rangle\rangle_x = 0 \quad \forall \texttt{v}_x \in \mathcal{V}_x, \quad \langle\langle \mathbf{1}_x, \texttt{GRADX p}\rangle\rangle_x = 0 \quad \forall \texttt{p} \in \mathcal{V}_\rho. \tag{36}$$

However, note that the third term in the right-hand side of (35) may cause a change in total momentum, just like the continuous term which it is an approximation of.

The zero integrals for the divergence (32), advection and gradient operators (36) are sufficient to guarantee mass and momentum conservation. They are easily met when discretization is carried out in practice, because of telescoping properties. Energy conservation introduces requirements which are investigated in the following sections. These requirements put additional conditions on the discretization methods.

4.1.3. Time derivative of the energy

Conservation of the integrated total energy $E_{kin} + E_{pot}$ is investigated by determining its time derivative. We will discuss both contributions individually.

Combining the discrete momentum equation (33) and the continuity equation (30) as in (5), the time derivative of the x-contribution of the integrated kinetic energy E_{kin} (28) becomes equal to

$$\begin{aligned}\frac{\partial E_{kin,x}}{\partial t} &= \frac{1}{2}\frac{\partial}{\partial t}\langle\langle \texttt{v}_x, (\texttt{S}_x\rho)\texttt{v}_x\rangle\rangle_x = \left\langle\left\langle \texttt{v}_x, \frac{\partial}{\partial t}\{(\texttt{S}_x\rho)\texttt{v}_x\}\right\rangle\right\rangle_x - \frac{1}{2}\left\langle\left\langle \texttt{v}_x, \texttt{v}_x\frac{\partial}{\partial t}(\texttt{S}_x\rho)\right\rangle\right\rangle_x \\ &\overset{(30),\,(33)}{=} -\langle\langle \texttt{v}_x, \texttt{GRADX p} + \hat{\rho}_x\ \texttt{GRADX}\ \phi\rangle\rangle_x - \left\langle\left\langle \texttt{v}_x, \left\{\texttt{ADV}_x - \frac{1}{2}\texttt{S}_x\ (\texttt{DIV q})\right\}\texttt{v}_x\right\rangle\right\rangle_x.\end{aligned} \tag{37}$$

Note that the kinetic energy is built from contributions defined in cell-faced velocity points, e.g. from $v_x \in \mathcal{V}_x$. Hence the density ρ and the contribution $\texttt{DIV q}$ from the continuity equation have to be interpolated towards the cell faces, based on their values in the cell centers where they are defined originally; compare (12) and (14). This explains the frequent appearance of the interpolation operator $\texttt{S}_x : \mathcal{V}_\rho \to \mathcal{V}_x$. The last line in (37) is the abstraction of the expression (A.8), see Appendix A.1, in case of the one-dimensional example of Section 3.

The time derivative of the integrated potential energy E_{pot} (29), summed over the density cells, is

$$\begin{aligned}\frac{\partial E_{pot}}{\partial t} &\overset{(29)}{=} \left\langle\left\langle \mathbf{1}_\rho, \frac{\partial}{\partial t}\{(\Phi+\phi)\ \rho - \texttt{p}\}\right\rangle\right\rangle_\rho = \left\langle\left\langle \mathbf{1}_\rho, (\Phi+\phi)\ \frac{\partial\rho}{\partial t} + \rho\frac{\partial\Phi}{\partial t} - \frac{\partial \texttt{p}}{\partial t}\right\rangle\right\rangle_\rho \overset{(4)}{=} \left\langle\left\langle \frac{\partial\rho}{\partial t}, \Phi+\phi\right\rangle\right\rangle_\rho \\ &\overset{(30)}{=} -\langle\langle \texttt{DIVX}\ \texttt{q}_x + \texttt{DIVY}\ \texttt{q}_y + \texttt{DIVZ}\ \texttt{q}_z, \Phi+\phi\rangle\rangle_\rho.\end{aligned} \tag{38}$$

In general, there is an exchange between kinetic and potential energy. However, the terms involved in this exchange can be identified. We will do so in the following two sections. A small complication is that the evolution of kinetic energy (37) is defined at cell faces, whereas the evolution of potential energy (38) is defined at cell centers. Therefore, in the one-dimensional example in Section 3 we started with the interpolation to cell faces, cf. (18) and (A.10).

4.1.4. Potential energy balance

Just like in the continuous case, certain terms in the change in potential energy (38) should cancel with certain terms in (37), the change of kinetic energy $E_{kin,x}$. Other terms in (38) should cancel with the changes of the kinetic energies $E_{kin,y}$ and $E_{kin,z}$. The remaining, convection related, terms in the changes of the kinetic energies should be zero, as is discussed in Section 4.1.5. The canceling terms in (38) and (37) are

$$\begin{aligned}\langle\langle \texttt{DIVX}\ \texttt{q}_x, \Phi+\phi\rangle\rangle_\rho + \langle\langle \texttt{v}_x, \texttt{GRADX p} + \hat{\rho}_x\ \texttt{GRADX}\ \phi\rangle\rangle_x &= \left\{\langle\langle \texttt{DIVX}\ \texttt{q}_x, \Phi\rangle\rangle_\rho + \langle\langle \texttt{v}_x, \texttt{GRADX p}\rangle\rangle_x\right\} \\ &\quad + \left\{\langle\langle \texttt{DIVX}\ \texttt{q}_x, \phi\rangle\rangle_\rho + \langle\langle \texttt{v}_x, \hat{\rho}_x\ \texttt{GRADX}\ \phi\rangle\rangle_x\right\}.\end{aligned} \tag{39}$$

Notice how in the last line the terms have been grouped in pairs that should cancel each other.

The discharge density $\hat{\rho}_x$ is defined analogously to (10). It is the solution of the product rule

$$\hat{\rho}_x\ \texttt{GRADX}\ \Phi = \texttt{GRADX p}. \tag{40}$$

The resulting density is always defined and the value in each x-grid point lies between the two density values in the grid points used for the partial derivative in that point, under the following conditions:

- Only a two-point stencil is used for the gradient,
- Φ is a monotonously increasing function of $\mathtt{p}$.

Note that under these conditions $\mathtt{GRADX}\ \Phi$ in (40) only vanishes when $\mathtt{GRADX\ p}$ vanishes.

Using the product rule (40), it is found that the potential energy balance is fulfilled when we choose the x-components $\mathtt{DIVX}$ of the discrete divergence and $\mathtt{q}_x$ of the discharge according to

$$\mathtt{q}_x \ := \ \hat{\rho}_x\ \mathtt{v}_x, \tag{41}$$

$$\langle\langle \mathtt{DIVX\ q}_x, \phi\rangle\rangle_\rho \ := \ -\langle\langle \mathtt{q}_x, \mathtt{GRADX}\ \phi\rangle\rangle_x \quad \forall \mathtt{q}_x \in \mathcal{V}_x,\ \ \phi \in \mathcal{V}_\rho. \tag{42}$$

E.g. the first term between the first pair of braces in (39) can be rewritten as

$$\langle\langle \mathtt{DIVX\ q}_x, \Phi\rangle\rangle_\rho \overset{(42)}{=} -\langle\langle \mathtt{q}_x, \mathtt{GRADX}\ \Phi\rangle\rangle_x \overset{(41)}{=} -\langle\langle \hat{\rho}_x\ \mathtt{v}_x, \mathtt{GRADX}\ \Phi\rangle\rangle_x = -\langle\langle \mathtt{v}_x, \hat{\rho}_x\ \mathtt{GRADX}\ \Phi\rangle\rangle_x \overset{(40)}{=} -\langle\langle \mathtt{v}_x, \mathtt{GRADX\ p}\rangle\rangle_x.$$

It obviously cancels the second term between the first pair of braces in (39). The terms between the second pair of braces in (39) cancel similarly. From (42), the discrete divergence can be calculated as soon as the discrete gradient is known. Remember, see (32), that the integral of the discrete divergence must be zero. Using (42) with $\phi \equiv \mathbf{1}_\rho$, this means that the discrete gradient of a constant field must be zero. This is a very reasonable property to require from the gradient.

4.1.5. Kinetic energy balance

The velocity-related terms in the change of kinetic energy (37) cancel when

$$\left\langle\left\langle \mathtt{v}_x, \left\{ \mathtt{ADV}_x - \frac{1}{2}\mathtt{S}_x(\mathtt{DIV\ q}) \right\} \mathtt{v}_x \right\rangle\right\rangle_x = 0 \quad \text{for all real-valued } \mathtt{v}_x \in \mathcal{V}_x, \tag{43}$$

and similar for the y- and z-directions. From Lemma 1 in Appendix B we learn that this requirement is equivalent to demanding that the operator

$$\mathtt{A}_x \ := \ \mathtt{ADV}_x - \frac{1}{2}\mathtt{S}_x(\mathtt{DIV\ q})\ \mathtt{I}_x \text{ is skew symmetric}, \tag{44}$$

where $\mathtt{I}_x$ is the unit operator in $\mathcal{V}_x$. Further, Lemma 2 shows that skew symmetry is a necessary condition for obtaining a divergence form in the corresponding energy.

Analytically, skew symmetry of the continuous equivalent of $\mathtt{A}$, the operator $\mathcal{A}$ given by $\mathcal{A}v := \nabla\cdot(\mathtt{q}v) - \frac{1}{2}(\nabla\cdot\mathtt{q})v$, is proven in Lemma 3 (Appendix B). We want this property to hold also for its discrete counterpart. This property was also identified as essential for energy conservation in [10,14,39].

Condition (44) means that the skew-symmetric part of the advection operator must compensate the change in the total momentum which is caused by the symmetric part. The symmetric part of the advection operator describes the change in (local) momentum due to changes in the density, and the skew-symmetric part describes the change in momentum due to changes in the velocity. It shows the close relationship between the advection operator and the continuity equation. Therefore, the advection operator will be constructed from the discrete operators $\mathtt{DIVX}$, $\mathtt{DIVY}$ and $\mathtt{DIVZ}$ used in the continuity equation. Before this can be done, some interpolations have to be introduced, because these operators do not work on the grid used for x-components (which is what $\mathtt{ADV}_x$ must work on); compare Fig. 5. Apart from the interpolation matrix $\mathtt{S}_x : \mathcal{V}_\rho \to \mathcal{V}_x$ (interpolation from scalar field to x-component), we need the new interpolations $\mathtt{S}_{xy} : \mathcal{V}_x \to \mathcal{V}_y$ and $\mathtt{S}_{xz} : \mathcal{V}_x \to \mathcal{V}_z$ (interpolation between vector components). These were not necessary in the one-dimensional example of Section 3, because in one dimension there is only one vector component.

Using the partial derivatives from the continuity equation and the interpolations, a preliminary approximation $\mathtt{ADV}_{x,0}$ for the advection operator is given by

$$\mathtt{ADV}_{x,0} \ := \ \mathtt{S}_x(\mathtt{DIVX}\ (\mathtt{q}_x\ \mathtt{I}_x) + \mathtt{DIVY}\ (\mathtt{q}_y\ \mathtt{S}_{xy}) + \mathtt{DIVZ}\ (\mathtt{q}_z\ \mathtt{S}_{xz})), \tag{45}$$

which is a discrete analog of $\nabla\cdot(\boldsymbol{q}v_x)$; compare (34). This is a sensible approximation for the advection operator, whose accuracy depends on the accuracies of the operators and interpolations used, but it does not fulfil the requirements (43) and (44).

Now, a suitable advection operator which satisfies the requirements is found by combining its skew-symmetric part and its diagonal, which can be determined from (44):

$$\mathtt{ADV}_x \ := \ \frac{1}{2}\mathtt{ADV}_{x,0} - \frac{1}{2}\mathtt{ADV}_{x,0}^T + \frac{1}{2}\mathtt{S}_x(\mathtt{DIV\ q})\mathtt{I}_x. \tag{46}$$

Observe that this expression corresponds with a discrete counterpart of an expression already used by e.g. Feiereisen et al. [7]:

$$\frac{1}{2}\nabla\cdot(\mathtt{q}\,v_x)+\frac{1}{2}(\mathtt{q}\cdot\nabla)v_x+\frac{1}{2}(\nabla\cdot\mathtt{q})v_x.$$

From its definition, it is clear that the choice (46) meets the skew-symmetry requirement (44). Only the momentum conserving property (36) remains to be checked. The integral of the preliminary approximation (45) is

$$\langle\langle\mathbf{1}_x,\mathtt{ADV}_{x,0}\,\mathtt{v}_x\rangle\rangle_x \overset{(45)}{=} \langle\langle\mathbf{1}_x,\mathtt{S}_x(\mathtt{DIVX}(\mathtt{q}_x\ \mathtt{I}_x)+\mathtt{DIVY}(\mathtt{q}_y\ \mathtt{S}_{xy})+\mathtt{DIVZ}(\mathtt{q}_z\ \mathtt{S}_{xz}))\mathtt{v}_x\rangle\rangle_x \overset{(27)}{=} \langle\langle\mathbf{1}_\rho,(\mathtt{DIVX}(\mathtt{q}_x\ \mathtt{I}_x)+\mathtt{DIVY}(\mathtt{q}_y\ \mathtt{S}_{xy})+\mathtt{DIVZ}(\mathtt{q}_z\ \mathtt{S}_{xz}))\mathtt{v}_x\rangle\rangle_\rho=\mathbf{0},$$

because each of the differential operators has zero integrals, cf. (32). This was to be expected since (45) is a discrete version of a divergence. Next, the integrals of the second and third term in the right-hand side of (46) are found to cancel each other. This follows by rewriting the integral of the second term as

$$\left\langle\left\langle\mathbf{1}_x,\mathtt{ADV}^{\,T}_{x,0}\mathtt{v}_x\right\rangle\right\rangle_x = \langle\langle\mathtt{ADV}_{x,0}\mathbf{1}_x,\mathtt{v}_x\rangle\rangle_x \overset{(45)}{=} \langle\langle\mathtt{S}_x(\mathtt{DIVX}(\mathtt{q}_x\ \mathtt{I}_x)+\mathtt{DIVY}(\mathtt{q}_y\ \mathtt{S}_{xy})+\mathtt{DIVZ}(\mathtt{q}_z\ \mathtt{S}_{xz}))\ \mathbf{1}_x,\mathtt{v}_x\rangle\rangle_x$$
$$= \langle\langle\mathtt{S}_x(\mathtt{DIVX}\ \mathtt{q}_x+\mathtt{DIVY}\ \mathtt{q}_y+\mathtt{DIVZ}\ \mathtt{q}_z),\mathtt{v}_x\rangle\rangle_x, \tag{47}$$

because $\mathtt{S}_{xy}\mathbf{1}_x=\mathbf{1}_y$, etc., for any reasonable interpolation operator. This expression clearly cancels the contribution of the third term in (46) which contains the diagonal of the advection operator. Hence the integral of the advection operator is zero and momentum conservation of the discrete formulation is guaranteed.

4.1.6. Relation with the 1D example from Section 3

To make the above presentation more concrete, it is now related to the one-dimensional example from Section 3. Especially, a comparison is made with the detailed formulas from Section 3.1.

Integrals and interpolation. In Section 4.1.1 abstract integration weight factors have been introduced. In the one-dimensional example these are given by $\mathtt{w}_\rho|_i=\Delta x_i$ and $\mathtt{w}_x|_{i+1/2}=\Delta x_{i+1/2}$. The interpolation of the density from the density points to the velocity points, i.e. $\rho_{i+1/2}:=\mathtt{S}_x\rho$, is chosen according to (12) and based on volume averaging. As demonstrated in (13), in this way the discrete total amount of mass is not dependent on the choice of the integration volumes. In a similar way, the potential energy has been interpolated in (18).

Divergence and gradient. The one-dimensional divergence and gradient operator follow from (8) and (11), respectively, as

$$\mathtt{DIV}\ \mathtt{q}|_i:=\frac{\mathtt{q}_{i+1/2}-\mathtt{q}_{i-1/2}}{\Delta x_i}\quad\text{and}\quad \mathtt{GRAD}\ \phi|_{i+1/2}:=\frac{\phi_{i+1}-\phi_i}{\Delta x_{i+1/2}}.$$

Note the integer and half-integer indices, indicating the staggered positions of these quantities; this corresponds with the function spaces $\mathcal{V}_\rho$ and $\mathcal{V}_x$, respectively. Further, note how the denominators cancel the integration weights in (42); hereto consider the respective inner products as defined in (24) and (25). Finally, recognize that (42) is a discrete summation-by-parts property, whence there is a close relation with the summation-by-parts method introduced in [27].

Energy balance. Section 4.1.4 shows that in the balance between potential and kinetic energy a crucial role is played by the discharge density $\hat{\rho}$; see (10). Its definition in (40) generalizes the discrete product rule (11).

Section 4.1.5 shows that also the form of the advection operator is essential. Its most important properties are the skew symmetry outside the diagonal (apart from the scaling with the control volume) and its entries on the diagonal (which have to cancel a contribution to $\mathtt{A}$ from the continuity equation). In (17) this has been incorporated in the following way. Firstly, the flux function $\mathtt{FV}$ is symmetric in $\mathtt{v}_{i+1/2}$ and $\mathtt{v}_{i-1/2}$, producing symmetry in the non-diagonal entries of the scaled coefficient matrix. Secondly, the choice of $\mathtt{q}_i$, defined in (15) as an average between the cell-faced discharges, ensures that the diagonal entry of $\mathtt{ADV}_x$ cancels the contribution from the continuity equation $\frac{1}{2}\mathtt{DIV}\ \mathtt{q}$. Both choices together make up for the required skew symmetry. Observe how in (A.9), from Appendix A.3, the corresponding terms match exactly. The abstract generalization hereof is presented above, compare e.g. the discussion around (46) and (47).

4.2. Construction of conservative methods using larger gradient stencils

The previous section describes the construction of MaMEC discretizations for regular structured grids. The discretizations were still subject to the restriction of a two-point stencil for the gradient operator. This restriction is problematic in certain cases:

- The order of accuracy can be at most second order when only two-point stencils are allowed for the gradient operator.
- In certain grid configurations, a two-point stencil is not possible. Consider for example the grid shown in Fig. 6. Full velocity vectors are approximated at every cell-center; scalars are approximated at the vertices. With cell-centers not 'half-way' vertices, standard two-point discretizations, e.g. for the gradient, are not adequate.

In this section therefore, a larger discretization stencil will be allowed for the gradient operator. Its usage is demonstrated on the unstructured grid from Fig. 6.

When a larger than two-point stencil is used for the gradient, the solution for Eq. (40) does not exist in grid points where the gradient of the potential function Φ is zero, but that of the pressure $\mathtt{p}$ is not. In other cases, the solution may have unrealistic values in it. To avoid these, a different approach must be taken for the potential energy balance. The problem is handled by factorizing the gradient operator. The (accurate, high-order) gradient $\mathtt{GRAD} : \mathcal{V}_\rho \to \mathcal{V}_3 := \mathcal{V}_x \times \mathcal{V}_y \times \mathcal{V}_z$ is written as the product of a two-point gradient $\mathtt{GRAD2} : \mathcal{V}_\rho \to \mathcal{V}_{aux}$ (which lives on an auxiliary grid, which may or may not be any of the other grids used in the discretization) and an interpolative factor $\mathtt{GRINT} : \mathcal{V}_{\mathrm{aux}} \to \mathcal{V}_3$, such that

$$\mathtt{GRAD} := \mathtt{GRINT}\ \mathtt{GRAD2}\ . \tag{48}$$

To construct such a factorization, the first step is to choose an auxiliary grid for which a discrete two-point gradient operator GRAD2 is 'easily' formed. Thereafter, the factor GRINT interpolates the gradient from the auxiliary grid to the velocity grid(s). To illustrate this procedure, an example is given below.

The procedure of creating a factorized gradient is rather difficult, and the results presented in this paper are not very accurate. More accurate factorized gradients can be obtained when so-called *Nedelec elements* are used, as described in [21]. We expect to present results with such methods in the near future.

Example. Fig. 6 shows a triangulation in which the discrete scalars, from the grid function space $\mathcal{V}_\rho$, are defined at vertices like *A* and *B*. The auxiliary grid consists of the points half-way the triangle edges, e.g. the point marked *AB*. Discrete velocity vectors, from $\mathcal{V}_3$, are defined at triangle centers, like the point *ABC*. Such a grid can be called 'staggered' in the sense that the scalars (pressure, density and potential energy) are defined on a different grid than the flow velocities. It can be called 'collocated' in the sense that all components of the flow velocities are defined on the same grid. The flow velocity's grid can be shown to have more grid points than the scalars' grid, so the only pressure field which causes no fluid motion is a constant pressure field.

To create a discrete gradient $\nabla\mathtt{p}$ in the triangle *ABC*, see the left-hand side of Fig. 6, first the operator GRAD2 is defined. It 'simply' is based on the difference between two neighbouring vertices. E.g. the choice for GRAD2 in the point *AB* is taken proportional to $\mathtt{p}_B - \mathtt{p}_A$. This yields a second-order approximation in the midpoints of the edges (e.g. *AB* in the auxiliary grid). Hereafter, these preliminary discrete gradient values are linearly combined into suitable values for the cell centers; this is the action of GRINT. The result becomes

$$\mathtt{w}_{ABC}\ \nabla\mathtt{p}|_{ABC} = \underline{\begin{pmatrix} y_{AB} - y_{ABC} \\ x_{ABC} - x_{AB} \end{pmatrix}}(\mathtt{p}_B - \mathtt{p}_A) + \begin{pmatrix} y_{ABC} - y_{AC} \\ x_{AC} - x_{ABC} \end{pmatrix}(\mathtt{p}_C - \mathtt{p}_A) + \begin{pmatrix} y_{ABC} - y_{BC} \\ x_{BC} - x_{ABC} \end{pmatrix}(\mathtt{p}_B - \mathtt{p}_C), \tag{49}$$

where $\mathtt{w}_{ABC}$ is the surface area of triangle *ABC*. This discrete gradient is also used in e.g. [21]. We will (and should, because of the relation between gradient and divergence) encounter the underlined vector, which is part of GRINT, again below in the mass transfer from *A* to *B*; see (54). The other vectors will be similarly found in the mass transfers from *A* to *C* and from *C* to *B*.

Like the operator GRAD in the product rule (40), the operator GRAD2, with its two-point stencil, can without problems be used in a product rule that defines an intermediate discharge density $\hat{\rho}$, similar to (10):

$$\hat{\rho}\ \mathtt{GRAD2}\ \Phi := \mathtt{GRAD2}\ \mathtt{p}. \tag{50}$$

Note that the discharge density is defined on the auxiliary grid, i.e. $\hat{\rho} \in \mathcal{V}_{aux}$. When the auxiliary grid does not coincide with the velocity grid, we have to interpolate the discharge density. Like for the gradient operator, the interpolation $\mathtt{GRINT} : \mathcal{V}_{aux} \to \mathcal{V}_3$ can be used: $\tilde{\rho} := \mathtt{GRINT}\ \hat{\rho}$. Now, similar to (41), a discharge $\mathtt{q}$ in the velocity grid points, i.e. $\mathtt{q} \in \mathcal{V}_3$, can be defined as

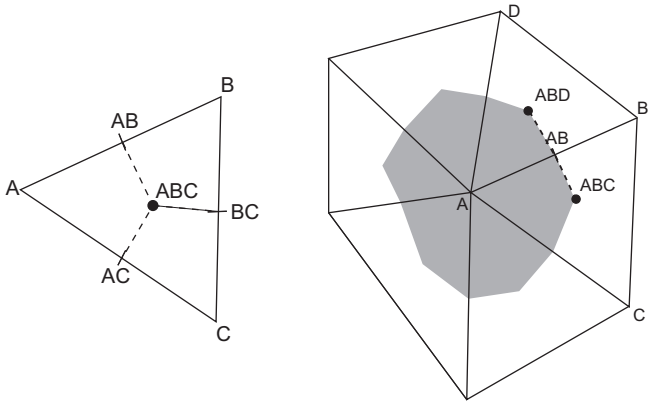

Fig. 6. A triangulation grid where scalars are defined at vertices and velocity vectors are defined in triangle centers. The left figure shows the stencil for the gradient in the center *ABC*; the right figure shows the control volume for the divergence around vertex *A*. Both discrete operators are closely related (see text).

$$\mathtt{q} := \tilde{\rho}\ \mathtt{v} = (\mathtt{GRINT}\ \hat{\rho})\ \mathtt{v}. \tag{51}$$

As discussed in Section 4.1.4, the relations (50) and (51) play a central role in the exchange between potential and kinetic energy. In particular the relation

$$\langle\langle \mathtt{DIV}\ \mathtt{q}, \Phi\rangle\rangle_\rho + \langle\langle \mathtt{v}, \mathtt{GRAD}\ \mathtt{p}\rangle\rangle_3 = 0$$

should hold; compare the terms between the first pair of brackets in (39). With the above definition of GRAD in (48), and using (51), this requirement can be rewritten as

$$\begin{aligned}&\langle\langle \mathtt{DIV}\ \mathtt{q}, \Phi\rangle\rangle_\rho + \langle\langle \mathtt{q}, \tilde{\rho}^{-1}\mathtt{GRINT}\ (\mathtt{GRAD2}\ \ \mathtt{p})\rangle\rangle_3 = 0\\ &\overset{(50)}{\Longleftrightarrow} \langle\langle \mathtt{q}, \mathtt{DIV}^{\,T}\Phi\rangle\rangle_3 + \langle\langle \mathtt{q}, \tilde{\rho}^{-1}\mathtt{GRINT}\ (\hat{\rho}\ \mathtt{GRAD2}\ \ \Phi)\rangle\rangle_3 = 0.\end{aligned}$$

The latter relation is clearly satisfied when the operator DIV is chosen according to

$$\mathtt{DIV}^{\,T} := -\tilde{\rho}^{-1}\ \mathtt{GRINT}(\hat{\rho}\ \mathtt{GRAD2}) \quad \text{i.e. } \mathtt{DIV} := -\mathtt{GRAD2}^{\,T}\ (\hat{\rho}\mathtt{GRINT}^{\,T}\tilde{\rho}^{-1}). \tag{52}$$

It should be verified that the divergence DIV, which has now become density-dependent, is still accurate: its density-dependence should be negligible up to the intended order of accuracy. In principle, this can be monitored by studying how fast the expression $\hat{\rho}\ \mathtt{GRINT}^{\,T}\tilde{\rho}^{-1} \to 1$ upon grid refinement.

The continuity equation, discretized on the scalar grid, becomes

$$\frac{\partial \rho}{\partial t} = -\mathtt{DIV}\ \mathtt{q} \overset{(51)}{=} -\mathtt{DIV}(\tilde{\rho}\ \mathtt{v}) \overset{(52)}{=} \mathtt{GRAD2}^{\,T}\ (\hat{\rho}\ \mathtt{GRINT}^{\,T}\mathtt{v}). \tag{53}$$

Note that this continuity equation (53) can be interpreted as a finite volume discretization: each column of $\mathtt{GRAD2}^T$ (which has only two non-zeros) moves an amount $\hat{\rho}\ \mathtt{GRINT}^{\,T}\mathtt{v}$ of mass-per-time from one grid point to another. This will next be illustrated with help of the above example.

Example (cont'd). Refer to the right-hand side drawing in Fig. 6. Its grey area represents the control volume for the conservation of mass. The mass-transfer flux $F_{A\to B}$ from point *A* to point *B* is chosen as

$$F_{A\to B} := \rho_{AB}\left(\underline{\begin{pmatrix} y_{AB}-y_{ABC}\\ x_{ABC}-x_{AB}\end{pmatrix}}\cdot\begin{pmatrix} u_{ABC}\\ v_{ABC}\end{pmatrix} + \begin{pmatrix} y_{ABD}-y_{AB}\\ x_{ABD}-x_{AB}\end{pmatrix}\cdot\begin{pmatrix} u_{ABD}\\ v_{ABD}\end{pmatrix}\right). \tag{54}$$

Here, ρ_{AB} defined in the point *AB*, which is part of the auxiliary grid, corresponds with the discharge density $\hat{\rho}$ as defined in (50). The linear combination of the velocity components corresponds with $\mathtt{GRINT}^T\mathtt{v}$. Together they make up the term between parentheses in (53). We have encountered the underlined vector already in the gradient for triangle *ABC*, see (49). The similar vector in the second term is found in the gradient for triangle *ABD*. This emphasizes the close relation between the discrete gradient and divergence operators, as expressed by (52).

With the above choice for DIV, the other term in the potential energy evolution that should cancel, see (39), can be written as

$$\langle\langle \mathtt{DIV}\ \mathtt{q}, \phi\rangle\rangle_\rho \overset{(51)}{=} \langle\langle \tilde{\rho}\ \mathtt{v}, \mathtt{DIV}^{\,T}\phi\rangle\rangle = \langle\langle \mathtt{v}, \tilde{\rho}\ \mathtt{DIV}^{\,T}\phi\rangle\rangle \overset{(52)}{=} -\langle\langle \mathtt{v}, \mathtt{GRINT}\ (\hat{\rho}\ \mathtt{GRAD2}\ \ \phi)\rangle\rangle.$$

This suggests to write the momentum equation as

$$\frac{\partial}{\partial t}\{(\mathtt{S}\rho)\mathtt{v}\} + \mathtt{ADV}\ \mathtt{v} + \mathtt{GRAD}\ \mathtt{p} + \mathtt{GRINT}\ (\hat{\rho}\ \mathtt{GRAD2}\ \ \phi) = 0.$$

The advective operator is chosen as in Section 4.1.5. Similar to (45) and (46), it is defined by

$$\mathtt{ADV} := \frac{1}{2}\mathtt{ADV}_{\,0} - \frac{1}{2}\mathtt{ADV}_{\,0}^{\,T} + \frac{1}{2}(\mathtt{S}\ \mathtt{DIV}\mathtt{q})\mathtt{I}_3 \quad \text{with } \mathtt{ADV}_{\,0} := \mathtt{S}\ \mathtt{DIV}(\mathtt{q}\mathtt{I}_3),$$

where $\mathtt{I}_3$ is the unit operator in $\mathcal{V}_3$. Observe that, as required in (44), the operator $\mathtt{A} := \mathtt{ADV} - \frac{1}{2}(\mathtt{S}\ \mathtt{DIV}\ \mathtt{q})\mathtt{I}_3$ is skew-symmetric. In the above definition, the interpolation operator $\mathtt{S} : \mathcal{V}_\rho \to \mathcal{V}_3$ between the scalar grid and the velocity grid appears. This operator is based on the integration weights in the respective inner products in order to satisfy the compatibility relation (13). Hereto consider in Fig. 6 the cross section of the velocity control volume (triangle *ABC*) and the mass control volume (the grey polygon). As *ABC* is the center of the triangle, this cross section covers exactly 1/3 of the surface area of triangle *ABC*. Thus the interpolation weights to be used in S are all equal to 1/3.

In this section, it was seen how large discretization stencils for the gradient operator can be accommodated in a MaMEC discretization. A factorization of the gradient operator is required in a two-point gradient and a smoother. The factorization is used to construct a density-dependent divergence, which is employed in the construction of the advection operator. The density-dependent divergence operator is also used in the continuity equation, but it does not really show itself in the formulation (53), because there it is written in terms of the velocity v and the discharge density $\hat{\rho}$.

5. Numerical example: 2d shallow water simulation

The conservation properties of the MaMEC discretization are illustrated in an example, where the shallow water equations are solved for a square domain. Again, very accurate time integration is used to restrict the effects of the time integration method. Focus is on the energy-conserving property of the MaMEC space discretization. In this way, computation times are higher than necessary; therefore no grid refinement studies have been performed. We postpone these until a compatible time-integration method has been implemented.

A simulation is carried out on a square domain of 5 by 5 meters with a flat bottom and periodic boundary conditions. The initial water level contains a circular elevation between 1.00 and 1.10 meters above the bottom, as is shown in the left picture of Fig. 7. The initial flow velocity is a constant flow field: $v = (2.5, 2.5)$ m/s. The advection carries the center of the water elevation to the top right of the domain, while the gravitational force spreads the elevation. The result is seen in the second picture of Fig. 7.

The calculation was carried out using three different discretizations:

- *MaMEC* the discretization described in Section 4.1.
  The results from Figs. 7 and 10 have been obtained this way. The results from the other discretizations are not shown because they look the same.
- *Finite Volume* a finite volume discretization, which conserves momentum and mass, but not energy.
  The main differences with the MaMEC discretization are that the intermediate density $\hat{\rho}$ is not used, and the advection operator does not have the required skew-symmetry property. It is very similar to the advection operator $\mathtt{ADV}_0$ presented in (45).
- *Lossy* a finite difference discretization, in which the advection operator does not conserve momentum.

This discretization is a direct approximation of the nonconservative momentum equation

$$\frac{\partial \mathrm{v}}{\partial t} + \mathrm{v} \cdot \nabla \mathrm{v} + \frac{1}{\rho} \nabla p = -\nabla \phi.$$

Obviously, this results in a 'nostalgic' method from the early days of CFD, when we still had to learn about the relevance of discrete conservation.

5.1. Rectilinear grid

The calculation is first done on a uniform rectilinear grid which has $120 \times 130 = 15600$ grid points. Fig. 7 shows the calculated water level fields.

The changes in the total discrete momentum and energy are compared among the calculations. In this 'lossy' method both the treatment of convection and of the discharges is different from the MaMEC treatment. Fig. 8 shows that the MaMEC discretization conserves momentum and energy, with an accuracy governed by the time integration scheme; with a sophisticated scheme [17,28] machine accuracy could even be obtained. In contrast, the finite volume discretization does not conserve energy and the 'lossy' discretization conserves only mass.

5.2. Triangular grid

For an application of a MaMEC discretization on an unstructured grid, the same calculation is done on a, rather irregular, triangulation shown in Fig. 9. The triangulation consists of 2008 triangles.

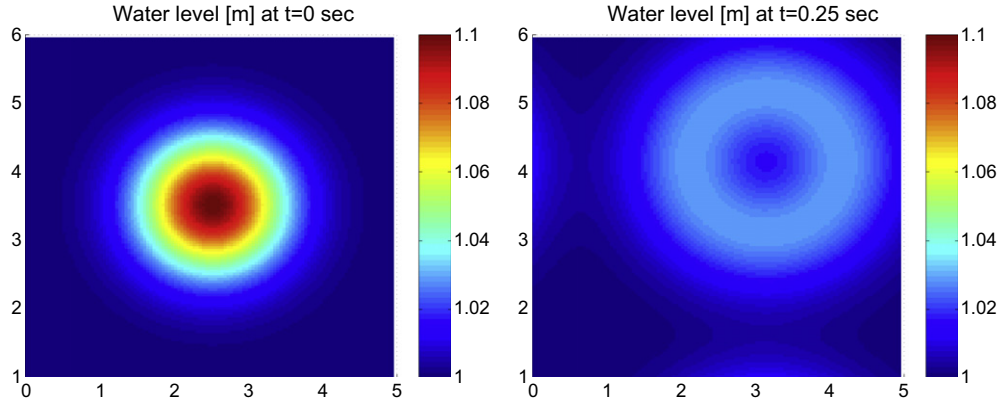

Fig. 7. Initial (left) and final (right) water levels calculated on the rectilinear grid.

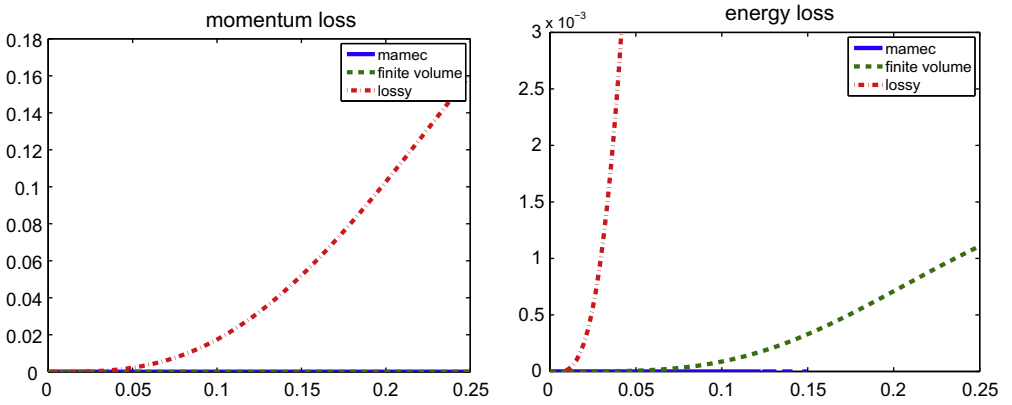

Fig. 8. Changes in momentum (left) and energy (right) in the rectilinear grid. Observe that only the MaMEC method conserves discrete energy.

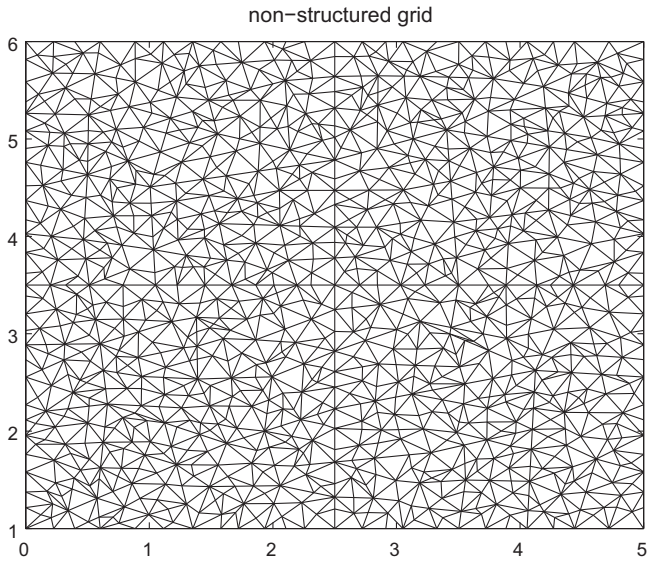

Fig. 9. The unstructured grid on which the calculation was repeated. Scalars (water levels, pressures, potential energy densities) are approximated at vertices, while vector components (velocities, mass fluxes) are approximated at cell centers.

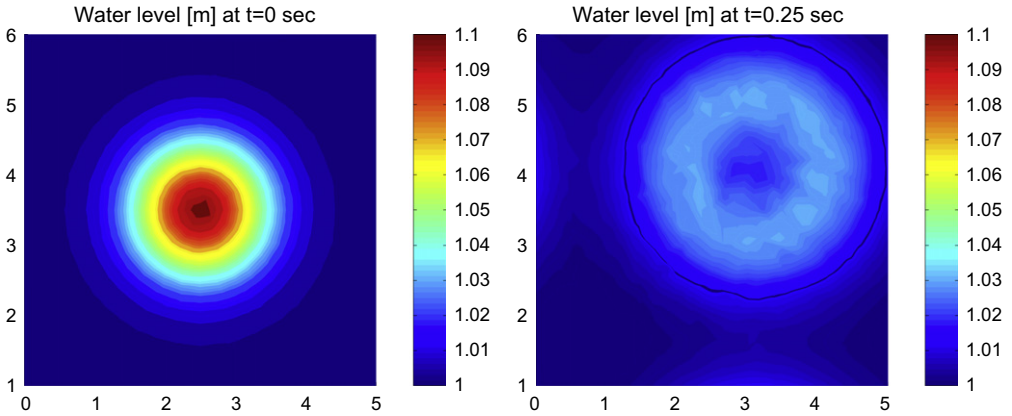

Fig. 10. Initial (left) and final (right) water levels calculated on a grid of 2008 triangles.

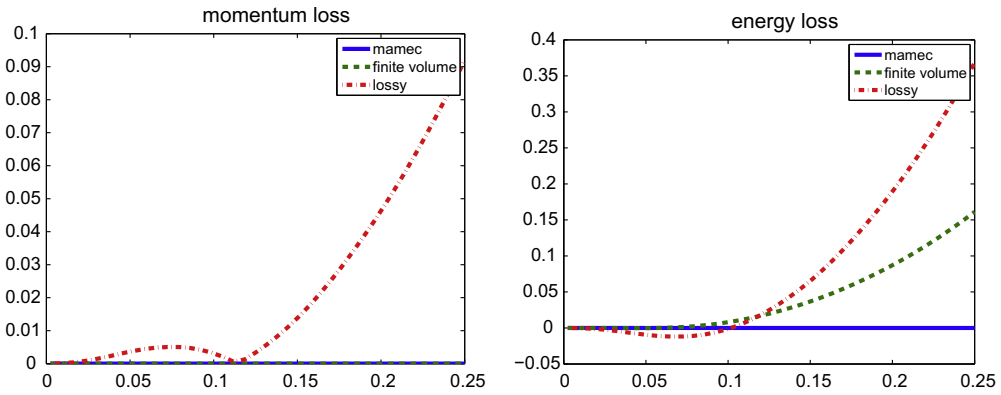

Fig. 11. Changes in momentum (left) and energy (right) on the grid consisting of triangles.

The density is approximated in the vertices, and both velocity components are approximated in the cell centers, as in Fig. 6. The discretizations on this grid are harder to construct than the ones in the rectilinear grid, because the gradient has a three-point stencil, as was pointed out in Section 4.2, so a factorization of the gradient is needed. This factorization was discussed with the help of Fig. 6.

Fig. 10 shows the calculated water level fields. The results are less smooth than those from the calculations on the rectilinear grid, but the latter grid had about eight times as many grid points: 15600 versus around 2000. The changes in total amounts of momentum and energy are presented similarly to the results of the rectilinear calculation in Fig. 11.

The MaMEC method combines the exact conservation of momentum and energy, and the application to general grids. These aspects may be essential in certain cases, as was seen in the one-dimensional example of Section 3, where the non-skew-symmetric method blows up.

6. Conclusions

This paper presents a method by which discretizations of the continuity and momentum equations can be designed for many classes of structured and unstructured grids. The discrete continuity and momentum equations can be combined into a discrete energy equation. The resulting 'MaMEC' discretizations conserve mass, momentum and energy, while typical finite volume methods only conserve mass and momentum. Two properties were found to be crucial:

- The contribution from convection has to be skew symmetric. This requires a close relation between momentum and continuity equation, in which the interpolation of fluxes and nodal values plays a subtle role.
- The product rule between pressure and energy transport potential has to hold in a discrete sense. This requires the definition of a separate discharge density.

The method was first illustrated with a one-dimensional example, in which the MaMEC method calculated qualitatively correct results, and a non-skew-symmetric finite volume method became unstable.

A second example was a simple shallow water calculation on a two-dimensional domain. Accurate results were obtained on a rectilinear grid, and on an unstructured grid. The conservation of mass, momentum and energy were checked and losses were negligible up to machine accuracy. The calculation on the unstructured grid requires the (non-unique) factorization of the gradient operator. A suitable factorization was proposed for the case of a triangulation grid.

Similar discretizations can be constructed for terms like viscosity, Coriolis forces, bottom friction, etcetera, which were not yet treated in the above discussion. These issues will be subject of further research.

Acknowledgements

The authors would like to thank the anonymous reviewers for their constructive remarks and suggestions which helped to improve the presentation.

Appendix A. Proofs

A.1. Proof of the discrete continuity equation (14)

In Section 3.1, a discrete continuity equation was presented for the (interpolated) cell-faced densities $\rho_{i+1/2}$. This continuity equation is derived in this section by filling in the interpolation formula (12) in the time derivative of the cell-faced density, and then replacing the time derivatives of the densities using the continuity equation (8).

$$\frac{\partial \rho_{i+1/2}}{\partial t} \overset{(12)}{=} \frac{1}{\Delta x_{i+1/2}} \left\{ \delta x_{i+1/4} \, \frac{\partial \rho_i}{\partial t} + \delta x_{i+3/4} \, \frac{\partial \rho_{i+1}}{\partial t} \right\} \tag{A.1}$$

$$\overset{(8)}{=} -\frac{1}{\Delta x_{i+1/2}} \left\{ \delta x_{i+1/4} \frac{\mathtt{q}_{i+1/2} - \mathtt{q}_{i-1/2}}{\Delta x_i} + \delta x_{i+3/4} \frac{\mathtt{q}_{i+3/2} - \mathtt{q}_{i+1/2}}{\Delta x_{i+1}} \right\}$$

$$= -\frac{1}{\Delta x_{i+1/2}} \left\{ \frac{\delta x_{i+3/4}}{\Delta x_{i+1}} \mathtt{q}_{i+3/2} + \left(\frac{\delta x_{i+1/4}}{\Delta x_i} - \frac{\delta x_{i+3/4}}{\Delta x_{i+1}} \right) \mathtt{q}_{i+1/2} - \frac{\delta x_{i+1/4}}{\Delta x_i} \mathtt{q}_{i-1/2} \right\}. \tag{A.2}$$

Next observe that

$$\frac{\delta x_{i+1/4}}{\Delta x_i} = \frac{\Delta x_i - \delta x_{i-1/4}}{\Delta x_i} = 1 - \frac{\delta x_{i-1/4}}{\Delta x_i} \quad \text{and similarly} \quad \frac{\delta x_{i+3/4}}{\Delta x_{i+1}} = 1 - \frac{\delta x_{i+5/4}}{\Delta x_{i+1}}. \tag{A.3}$$

Herewith (A.2) can be rewritten as

$$\frac{\partial \rho_{i+1/2}}{\partial t} \overset{(A.3)}{=} -\frac{1}{\Delta x_{i+1/2}} \left\{ \frac{\delta x_{i+3/4}}{\Delta x_{i+1}} \mathtt{q}_{i+3/2} - \left(\frac{\delta x_{i-1/4}}{\Delta x_i} - \frac{\delta x_{i+5/4}}{\Delta x_{i+1}} \right) \mathtt{q}_{i+1/2} - \frac{\delta x_{i+1/4}}{\Delta x_i} \mathtt{q}_{i-1/2} \right\}$$

$$= -\frac{1}{\Delta x_{i+1/2}} \left\{ \frac{\delta x_{i+3/4} \ \mathtt{q}_{i+3/2} + \delta x_{i+5/4} \ \mathtt{q}_{i+1/2}}{\Delta x_{i+1}} - \frac{\delta x_{i-1/4} \ \mathtt{q}_{i+1/2} + \delta x_{i+1/4} \ \mathtt{q}_{i-1/2}}{\Delta x_i} \right\} \tag{A.4}$$

$$\overset{(15)}{=} -\frac{\mathtt{q}_{i+1} - \mathtt{q}_i}{\Delta x_{i+1/2}}.$$

Thus (14) has been proven. Observe how the interpolation choice in (12), which results in (A.1), straightforwardly leads to (A.4). This then defines the interpolation strategy for $\mathtt{q}$ as introduced in (15).

A.2. Proof of the continuous energy equation (6)

This section steps through the derivation of the continous energy equation (6), because similar steps have to be made in the derivation of the discrete energy equation (19) (the latter will be carried out in Appendix A.3). It is important to recognize which terms are to be combined, and hence have to be compatible. E.g. from the thermodynamics side, the time derivative of the pressure can be expressed in terms of the density and the time derivative of the potential function Φ according to (4):

$$\frac{\partial p}{\partial t} = \rho \frac{\partial \Phi}{\partial t} \quad \text{hence} \quad \frac{\partial \rho e_{pot}}{\partial t} = \frac{\partial}{\partial t}\{\rho(\Phi + \phi) - p\} = (\Phi + \phi)\frac{\partial \rho}{\partial t}. \tag{A.5}$$

Following (5), it is noted that

$$\frac{\partial \rho e_{kin}}{\partial t} = \frac{\partial}{\partial t}\left(\frac{1}{2}\rho|\boldsymbol{v}|^2\right) = \boldsymbol{v} \cdot \frac{\partial \rho|\boldsymbol{v}|}{\partial t} - \frac{1}{2}|\boldsymbol{v}|^2 \frac{\partial \rho}{\partial t}. \tag{A.6}$$

Making use of (A.6) and (A.5), we can write

$$\frac{\partial \rho e}{\partial t} = \frac{\partial \rho e_{kin}}{\partial t} + \frac{\partial \rho e_{pot}}{\partial t} = |\boldsymbol{v}| \cdot \frac{\partial \rho|\boldsymbol{v}|}{\partial t} - \frac{1}{2}|\boldsymbol{v}|^2 \frac{\partial \rho}{\partial t} + (\Phi + \phi)\frac{\partial \rho}{\partial t}$$

$$= -|\boldsymbol{v}| \cdot \{\nabla \cdot (\mathbf{q} \otimes |\boldsymbol{v}|) + \nabla p + \rho \nabla \phi\} + \frac{1}{2}|\boldsymbol{v}|^2 \nabla \cdot \mathbf{q} - (\Phi + \phi)\nabla \cdot \mathbf{q}$$

$$\overset{(4)}{=} -\frac{1}{2}\nabla \cdot \mathbf{q}|\boldsymbol{v}|^2 - |\boldsymbol{v}| \cdot \rho \nabla(\Phi + \phi) - (\Phi + \phi)\nabla \cdot \mathbf{q}$$

$$= -\frac{1}{2}\nabla \cdot \mathbf{q}|\boldsymbol{v}|^2 - \nabla \cdot \mathbf{q}(\Phi + \phi) = -\nabla \cdot \mathbf{q} \left(\frac{1}{2}|\boldsymbol{v}|^2 + \Phi + \phi\right). \tag{A.7}$$

Hereafter (6) follows. Note in the last lines of (A.7) how terms from different origin are being combined into a divergence form. Such a combination also has to work in the discrete counterpart, to be discussed below.

A.3. Proof of the discrete energy equation (19)

In Section 3.1, a finite volume form is given for the discrete energy equation. This conservative energy equation is derived in this section Appendix A.2 gives a continuous counterpart).

Time differentiation of the kinetic energy contribution in (18) leads to

$$\left.\frac{\partial \rho e_{kin}}{\partial t}\right|_{i+1/2} \overset{(A.6)}{=} \mathrm{v}_{i+1/2}\frac{\partial}{\partial t}(\rho_{i+1/2}v_{i+1/2}) - \frac{1}{2}\mathrm{v}_{i+1/2}^2\frac{\partial \rho_{i+1/2}}{\partial t}$$

$$\overset{(14),(16)}{=} -\mathrm{v}_{i+1/2}\left\{\frac{\mathrm{FV}_{i+1} - \mathrm{FV}_i}{\Delta x_{i+1/2}} + \hat{\rho}_{i+1/2}\frac{\phi_{i+1}-\phi_i}{\Delta x_{i+1/2}}\right\} + \frac{1}{2}\mathrm{v}_{i+1/2}^2\frac{\mathrm{q}_{i+1}-\mathrm{q}_i}{\Delta x_{i+1/2}}$$

$$\overset{(9),(17)}{=} \underbrace{-\frac{1}{2}\mathrm{v}_{i+1/2}\frac{\mathrm{q}_{i+1}(\mathrm{v}_{i+3/2}+\mathrm{v}_{i+1/2}) - \mathrm{q}_i(\mathrm{v}_{i+1/2}+\mathrm{v}_{i-1/2})}{\Delta x_{i+1/2}}}_{\text{convection}}$$

$$\underbrace{-\mathrm{v}_{i+1/2}\frac{p_{i+1}-p_i}{\Delta x_{i+1/2}}}_{\text{pressure gradient}} \underbrace{-\mathrm{q}_{i+1/2}\frac{\phi_{i+1}-\phi_i}{\Delta x_{i+1/2}}}_{\text{thermodynamics}} \underbrace{+\frac{1}{2}\mathrm{v}_{i+1/2}^2\frac{\mathrm{q}_{i+1}-\mathrm{q}_i}{\Delta x_{i+1/2}}}_{\text{continuity}} \tag{A.8}$$

$$\overset{(11)}{=} -\frac{1}{2}\frac{\mathrm{q}_{i+1}\mathrm{v}_{i+3/2}\mathrm{v}_{i+1/2} - \mathrm{q}_i\mathrm{v}_{i+1/2}\mathrm{v}_{i-1/2}}{\Delta x_{i+1/2}} - \mathrm{q}_{i+1/2}\frac{(\Phi_{i+1}+\phi_{i+1}) - (\Phi_i+\phi_i)}{\Delta x_{i+1/2}}. \tag{A.9}$$

In the step between (A.8) and (A.9) several terms have to cancel each other. Firstly, notice how the contribution from the continuity equation cancels against the diagonal contribution from the convective term. This is part of the skew-symmetry requirement (44). Also, the coefficients of the neighboring points $\mathrm{v}_{i-1/2}$ and $\mathrm{v}_{i+1/2}$ have to be related. This is taken care of by the (constant) averaging in (17). Finally, the contribution from the pressure gradient and the thermodynamic term have to fit together. This is taken care of by the discrete product rule (11).

Differentiation of the potential energy contribution in (18), using (A.5), then results in

$$\left.\frac{\partial \rho e_{pot}}{\partial t}\right|_{i+1/2} \overset{(A.5)}{=} \frac{\delta x_{i+1/4}}{\Delta x_{i+1/2}}(\Phi_i+\phi_i)\frac{\partial \rho_i}{\partial t} + \frac{\delta x_{i+3/4}}{\Delta x_{i+1/2}}(\Phi_{i+1}+\phi_{i+1})\frac{\partial \rho_{i+1}}{\partial t} \overset{(8)}{=} -\frac{\delta x_{i+1/4}}{\Delta x_{i+1/2}}(\Phi_i+\phi_i)\frac{\mathrm{q}_{i+1/2}-\mathrm{q}_{i-1/2}}{\Delta x_i}$$
$$-\frac{\delta x_{i+3/4}}{\Delta x_{i+1/2}}(\Phi_{i+1}+\phi_{i+1})\frac{\mathrm{q}_{i+3/2}-\mathrm{q}_{i+1/2}}{\Delta x_{i+1}}. \tag{A.10}$$

Gathering the terms in (A.9) and (A.10) that contain $\Phi_i + \phi_i$ leads to

$$\mathrm{q}_{i+1/2}\frac{(\Phi_i+\phi_i)}{\Delta x_{i+1/2}} - \frac{1}{\Delta x_{i+1/2}}\frac{\delta x_{i+1/4}}{\Delta x_i}(\Phi_i+\phi_i)(\mathrm{q}_{i+1/2}-\mathrm{q}_{i-1/2}) = \frac{1}{\Delta x_{i+1/2}}(\Phi_i+\phi_i)$$
$$\left\{\left(1-\frac{\delta x_{i+1/4}}{\Delta x_i}\right)\mathrm{q}_{i+1/2} + \frac{\delta x_{i+1/4}}{\Delta x_i}\mathrm{q}_{i-1/2}\right\} \overset{(A.3)}{=} \frac{1}{\Delta x_{i+1/2}}(\Phi_i+\phi_i)\left\{\frac{\delta x_{i-1/4}}{\Delta x_i}\mathrm{q}_{i+1/2} + \frac{\delta x_{i+1/4}}{\Delta x_i}\mathrm{q}_{i-1/2}\right\} \overset{(15)}{=} \frac{1}{\Delta x_{i+1/2}}(\Phi_i+\phi_i)\ \mathrm{q}_i; \tag{A.11}$$

a similar relation holds for the terms that contain $\Phi_{i+1} + \phi_{i+1}$.

Combination of the full Eqs. (A.9) and (A.10) herewith results in a discrete analogue of (A.7)

$$\left.\frac{\partial}{\partial t}\rho(e_{kin}+e_{pot})\right|_{i+1/2} = -\frac{1}{2}\frac{\mathrm{q}_{i+1}\mathrm{v}_{i+3/2}\mathrm{v}_{i+1/2} - \mathrm{q}_i\mathrm{v}_{i+1/2}\mathrm{v}_{i-1/2}}{\Delta x_{i+1/2}} - \frac{\mathrm{q}_{i+1}(\Phi_{i+1}+\phi_{i+1}) - \mathrm{q}_i(\Phi_i+\phi_i)}{\Delta x_{i+1/2}},$$

which can be written in the divergence form (19).

Appendix B. Skew symmetry

Without being too precise about the function spaces involved, only compact support is essential to get rid of boundary terms, a real-valued operator $\mathcal{A}$ is called skew symmetric when

$$\int_\Omega v\ (\mathcal{A}w)\,\mathrm{d}\Omega = -\int_\Omega (\mathcal{A}v)\ w\,\mathrm{d}\Omega \quad \forall \text{ real } v,w. \tag{B.1}$$

The following three lemmas are useful to identify skew-symmetric operators.

Lemma 1. *A real-valued operator $\mathcal{A}$ is skew symmetric if and only if $\int_\Omega v\ \mathcal{A}v\,d\Omega = 0$ for all real-valued v.*

Proof. $\Rightarrow$: First, let A be skew symmetric. Then we have

$$\int_\Omega v\ (\mathcal{A}v)\,\mathrm{d}\Omega \overset{(B.1)}{=} -\int_\Omega (\mathcal{A}v)\ v\,\mathrm{d}\Omega = -\int_\Omega v\ (\mathcal{A}v)\,\mathrm{d}\Omega.$$

Hence both sides of the equality have to vanish.

$\Leftarrow$: For any real v, w we have

$$0 = \int_\Omega (v+w)\ \mathcal{A}(v+w)\,\mathrm{d}\Omega = \int_\Omega v\ (\mathcal{A}v)\,\mathrm{d}\Omega + \int_\Omega w\ (\mathcal{A}v)\,\mathrm{d}\Omega + \int_\Omega v\ (\mathcal{A}w)\,\mathrm{d}\Omega + \int_\Omega w\ (\mathcal{A}w)\,\mathrm{d}\Omega.$$

As inner products involving the same real function on both sides of the operator vanish, we conclude

$$\int_\Omega v\ (\mathcal{A}w)\,\mathrm{d}\Omega = -\int_\Omega w\ (\mathcal{A}v)\,\mathrm{d}\Omega \text{ for all real } v,\ w.$$

Hence A is skew symmetric. □

The next theorem relates skew symmetry of an operator to the conservation form of the related 'energy'.

Lemma 2. *Let the real-valued operator $\mathcal{A}$ be such that for all real v there exists a function $\mathcal{B}_v$ that satisfies $v\ \mathcal{A}v = \nabla\cdot B_v$,i.e. the energy-related expression $v\ \mathcal{A}v$ is of divergence type. Then $\mathcal{A}$ is skew symmetric.*

Proof. For all v we can write

$$\int_\Omega v\ \mathcal{A}v\,\mathrm{d}\Omega = \int_\Omega \nabla\cdot\mathcal{B}_v\,\mathrm{d}\Omega \overset{(a)}{=} \int_{\partial\Omega} n\cdot\mathcal{B}_v\,\mathrm{d}\partial\Omega = 0,$$

because all functions considered have compact support, whence after application of Gauss' divergence theorem, step (a), all terms at the boundary $\partial\Omega$ (with normal n) vanish. According to Lemma 1, the operator $\mathcal{A}$ is skew symmetric. □

Lemma 3. *The operator $\mathcal{A}$ given by*

$$\mathcal{A}v \ :=\ \nabla\cdot(\mathtt{q}v) - \frac{1}{2}(\nabla\cdot\mathtt{q})v$$

is skew symmetric.

Proof. Observe that

$$v\ \mathcal{A}v = \frac{1}{2}\nabla\cdot(\mathtt{q}vv),$$

after which Lemma 2 gives the desired property. □

References

[1] A. Arakawa, Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow, J. Comput. Phys. 1 (1966) 119–143.
[2] L. Bonaventura, L. Kornblueh, T. Heinze, P. Ripodas, A semi-implicit method conserving mass and potential vorticity for the shallow water equations on the sphere, Int. J. Numer. Methods Fluids 47 (2005) 863–869.
[3] E.J. Caramana, D.E. Burton, M.J. Shashkov, P.P. Whalen, The construction of compatible hydrodynamics algorithms utilizing conservation of total energy, J. Comput. Phys. 146 (1998) 227–262.
[4] M. Dröge, R. Verstappen, A new symmetry-preserving Cartesian-grid method for computing flow past arbitrarily shaped objects, Int. J. Numer. Methods Fluids 47 (2005) 979–985.
[5] F. Ducros, F. Laporte, T. Souléres, V. Guinot, P. Moinat, B. Caruelle, High-order fluxes for conservative skew-symmetric-like schemes in structured meshes: application to compressible flows, J. Comput. Phys. 161 (2000) 114–139.
[6] S. Dubinkina, J. Frank, Statistical relevance of vorticity conservation in the Hamiltonian particle-mesh method, J. Comput. Phys. 229 (2010) 2634–2648.
[7] W.J. Feiereisen, W.C. Reynolds, J.H. Ferziger, Numerical simulation of a compressible, homogeneous, turbulent shear flow. Report TF-13, Thermosciences Division, Mechanical Engineering, Stanford University, 1981.
[8] F.N. Felten, T.S. Lund, Kinetic energy conservation issues associated with the collocated mesh scheme for incompressible flow, J. Comput. Phys. 215 (2006) 465–484.
[9] J. Frank, S. Reich, Conservation properties of smoothed particle hydrodynamics applied to the shallow water equations, BIT Numer. Methods 43 (2003) 40–54.
[10] A. Jameson, The construction of discretely conservative finite volume schemes that also globally conserve energy or entropy, J. Sci. Comput. 234 (2008) 152–187.
[11] S. Kang, H. Pitsch, A robust scheme for compressible flows with a minimum artificial stabilization. Annual Research Briefs 2009, Center for Turbulent Research, Stanford University, 2009, pp. 353–365.
[12] B. Karasözen, V.G. Tsybulin, Cosymmetry preserving finite-difference methods for convection equations in a porous medium, Appl. Numer. Math. 55 (2005) 69–82.
[13] K.M.T. Kleefsman, G. Fekken, A.E.P. Veldman, B. Iwanowski, B. Buchner, A Volume-of-Fluid based simulation method for wave impact problems, J. Comput. Phys. 206 (2005) 363–393.
[14] J.C. Kok, A high-order low-dispersion symmetry-preserving finite-volume method for compressible flow on curvilinear grids, J. Comput. Phys. 228 (2009) 6811–6832.
[15] J.-G. Liu, W.-C. Wang, Energy and helicity preserving schemes for hydro- and magnetohydro-dynamics flows with symmetry, J. Comput. Phys. 200 (2004) 8–33.
[16] F. Mesinger, A. Arakawa, Numerical methods used in atmospheric models, GARP Publ. Ser. 1 (17) (1977).
[17] Y. Morinishi, Skew-symmetric form of convective terms and fully conservative finite difference schemes for variable density low-Mach number flows, J. Comput. Phys. 229 (2010) 276–300.
[18] Y. Morinishi, T.S. Lund, O.V. Vasilyev, P. Moin, Fully conservative finite difference schemes for incompressible flow, J. Comput. Phys. 143 (1998) 90–124.
[19] B. Perot, Conservation properties of unstructured staggered mesh schemes, J. Comput. Phys. 159 (2000) 58–89.
[20] J. Blair Perot, Discrete conservation properties of unstructured mesh schemes, Ann. Rev. Fluid Mech. 43 (2011) 299–318.
[21] B. Perot, V. Subramanian, Discrete calculus methods for diffusion, J. Comput. Phys. 224 (2007) 59–81.

[22] S.A. Piacsek, G.P. Williams, Conservation properties of convection difference schemes, J. Comput. Phys. 6 (1970) 392–405.
[23] M. van Reeuwijk, A mimetic mass, momentum and energy conserving discretization for the shallow water equations, Comput. Fluids 46 (2011) 411–416.
[24] Todd D. Ringler, David A. Randall, A potential enstrophy and energy conserving numerical scheme for solution of the shallow water equations on a geodesic grid, Mon. Weath. Rev. 130 (2002) 1397–1410.
[25] S. Steinberg, A discrete calculus with applications of high-order discretizations to boundary-value problems, Comput. Methods Appl. Math. 4 (2004) 228–261.
[26] G.S. Stelling, S.P.A. Duinmeijer, A staggered conservative scheme for every Froude number in rapidly varied shallow water flows, Int. J. Numer. Methods Fluids 43 (2003) 1329–1354.
[27] B. Strand, Summation by parts for finite difference approximations for d/dx, J. Comput. Phys. 110 (1994) 47–67.
[28] P.K. Subbareddy, G.V. Candler, A fully discrete, kinetic energy consistent finite-volume scheme for compressible flows, J. Comput. Phys. 228 (2009) 1347–1384.
[29] V. Subramanian, B. Perot, Higher-order mimetic methods for unstructured meshes, J. Comput. Phys. 219 (2006) 68–85.
[30] F.X. Trias, R.W.C.P. Verstappen, A. Gorobets, M. Soria, A. Oliva, Parameter-free symmetry-preserving regularization modeling of a turbulent differentially heated cavity, Comput. Fluids 39 (2010) 1815–1831.
[31] O.V. Vasilyev, High order difference schemes on non-uniform meshes with good conservation properties, J. Comput. Phys. 157 (2000) 746–761.
[32] A.E.P. Veldman, K.-W. Lam, Symmetry-preserving upwind discretization of convection on non-uniform grids, Appl. Numer. Math. 58 (2008) 1881–1891.
[33] A.E.P. Veldman, K. Rinzema, Playing with nonuniform grids, J. Eng. Math. 26 (1991) 119–130.
[34] A.E.P. Veldman, J. Gerrits, R. Luppes, J.A. Helder, J.P.B. Vreeburg, The numerical simulation of liquid sloshing on board spacecraft, J. Comput. Phys. 224 (2007) 82–99.
[35] R. Verstappen, On restraining the production of small scales of motion in a turbulent channel flow, Comput. Fluids 37 (2008) 887–897.
[36] R.W.C.P. Verstappen, A.E.P. Veldman, Direct numerical simulation of turbulence at lesser costs, J. Eng. Math. 32 (1997) 143–159.
[37] R.W.C.P. Verstappen, A.E.P. Veldman, Spectro-consistent discretization: a challenge to RANS and LES, J. Eng. Math. 34 (1998) 163–179.
[38] R.W.C.P. Verstappen, A.E.P. Veldman, Symmetry-preserving discretization of turbulent flow, J. Comput. Phys. 187 (2003) 343–368.
[39] R. Wemmenhove. Numerical simulation of two-phase flow in offshore environments. PhD Thesis, University of Groningen, The Netherlands, 2008.